

# ASIMETRÍA, PERSISTENCIA Y NO LINEALIDAD DE LA TASA DE DESEMPLEO ESPAÑOL

**José María Casado**

y

**F. Javier Trívez**

(Universidad de Zaragoza)

**Abstract:** La asimetría o naturaleza contracíclica y su influencia en la persistencia de la tasa registrada de desempleo, es un tema clásico de análisis de la teoría económica, que no ha sido atendido en los estudios que del desempleo español se han realizado, centrándose en demostrar su memoria larga y elaborando modelos macroeconómicos con vectores autorregresivos en los que la variable desempleo se supone no estacionaria y cointegrada. Su consideración, en este artículo, con modelos autorregresivos de transición suave (STAR), que detectan los diferentes regímenes existentes y la velocidad de cambio, no sólo nos permite una mejora en la especificación y predicción a corto plazo de la variable, sino que nos conducen a suponer al igual que Skalin y Teräsvirta (2002) un comportamiento del desempleo localmente no-estacionario en un modelo globalmente estable.

**Key Words:** Persistencia, asimetría, no-linealidad, modelos autorregresivos de transición suave, serie temporal, desempleo.

**CLASIFICACIÓN JEL:** C22, C52, E24

## **Address:**

Departamento de Análisis Económico  
Facultad Ciencias Económicas  
Gran Vía, 2-4  
Fax: 34 976761996  
Email: [fjtrivez@unizar.es](mailto:fjtrivez@unizar.es)

**Acknowledgements:** Esta investigación ha sido financiada por el Ministerio de Ciencia y Tecnología y FEDER bajo el proyecto SEC2002-02350 del Plan Nacional de Investigación Científica, Desarrollo e Innovación Tecnológica. Los autores desean agradecer al profesor Dick van Dijk sus sugerencias y útiles comentarios.

# 1. INTRODUCCIÓN

Las características del desempleo más importantes a la hora de analizar su evolución son, por una parte, su larga persistencia en el tiempo, y por otra, su marcada naturaleza contracíclica, mostrando un crecimiento acelerado en las etapas de recesión y una disminución atenuada en las de expansión.

La persistencia o permanencia del empleo en el nivel alcanzado es lo que tradicionalmente se ha llamado histéresis (Lindbeck y Snower, 1985; Blanchard y Summers, 1987). En términos econométricos la histéresis en el desempleo parece asumir memoria larga, en el sentido de raíces unitarias, y entiende como medida de la persistencia la suma de los coeficientes en un proceso autorregresivo. Una popular línea de investigación ha sido demostrar el fenómeno de la histéresis o larga persistencia en la tasa de desempleo chequeando la hipótesis nula de raíz unitaria, y midiendo la dimensión de la memoria a través de los coeficientes en un proceso ARMA (ver Sachs, 1986 y Blanchard y Summers, 1990).

En cambio estos estudios no tienen en cuenta características propias de la tasa del desempleo en la construcción de los modelos, que aconsejan la estimación por procedimientos no lineales y que pueden afectar significativamente a la hora de medir la persistencia de shocks puntuales. Bianchi y Zoega (1998) y Skalin y Teräsvirta (2002) asumen incluso que si esas peculiaridades son tenidas en cuenta puede afirmarse que la tasa de desempleo es globalmente estacionaria pero posiblemente no lineal y localmente no estacionaria.

Entre las razones que justifican la no linealidad se encuentra la clara naturaleza contracíclica del desempleo, incrementándose más rápidamente en etapas de recesión que lo que disminuye en los periodos de expansión económica. Este resultado es consistente con las teorías de costes asimétricos de contratación y despido de Bentolila y Bertola (1990) así como con la insider-outsider de Lindbeck y Snower (1988) que argumentan que en épocas de expansión económica el núcleo fijo de trabajadores presionará para trasladar a salarios más elevados los aumentos en la demanda agregada, en vez de incrementar el nivel de empleo. Esta habilidad de los insiders para trasladar a salarios los incrementos de la demanda está en clara conexión con el concepto de persistencia e histéresis, ya que shocks agregados pronunciados que desplacen la demanda de trabajo pueden tener efectos persistentes sobre los salarios y el empleo, siendo el efecto de persistencia mayor en los países con costes de rotación elevados y sindicatos fuertes.

Persistencia y asimetría, histéresis e *insiders-outsider* son conceptos y teorías conexas que intentan explicar las causas de las altas tasas de paro que existen en Europa y que especialmente en España parecen que se han hecho endémicas, sin que ni la moderación salarial ni las políticas expansivas logren reducirlas de una forma significativa.

Son estas características tan peculiares, las que hacen del desempleo, una variable que debe ser estudiada siguiendo modelos de serie temporal no lineal. Numerosos son los trabajos que se han desarrollado en este sentido, destacando como pionero el trabajo de Neftçi (1984) en el que utiliza la serie de desempleo como un indicador del ciclo

económico analizando las asimetrías mediante una transición de Markov de dos estados. Parker y Rothman (1997), Rohtman (1998), Montgomery et al (1998), y van Dijk, Teräsvirta y Franses (2002) analizan el desempleo de Estados Unidos con distintos modelos no lineales. Otros autores como Hansen (1997) han desarrollado la inferencia estadística de los parámetros de truncamiento en modelos autorregresivos de umbral (TAR). Bianchi y Zoega (1998) retoman el estudio de los equilibrios múltiples y analiza su existencia para los quince países de la OCDE usando una variante del modelo de regresión de cambio con transición tipo Markov introducido por Lindgren (1978). Koop y Potter (1999) desarrollan la transformación logística del desempleo aplicando técnicas bayesiana sobre los modelos TAR. Bränäs y Ohlsson (1999) generan fórmulas de agregación temporal para modelos simétricos autorregresivos de media móvil (ARasMA) aplicándolo al desempleo mensual y trimestral sueco. Por último citar los trabajos de Caner y Hansen (2001) y la modelización de las asimetrías de la tasa de desempleo para algunos países de la OCDE que realizan Skalin y Teräsvirta (2002).

A pesar de que muchas teorías económicas recomiendan tener en cuenta la naturaleza asimétrica del desempleo, estas no han sido muy atendidas en los diferentes estudios que del mercado de trabajo español se han realizado. Los modelos macroeconómicos tradicionalmente planteados son vectores autorregresivos en los que la variable desempleo se supone no estacionaria y cointegrada con el resto de variables macroeconómicas. En este artículo, tratamos de demostrar que la tasa de paro española debe modelizarse teniendo en cuenta sus asimetrías bajo un modelo autorregresivo de transición suave (STAR) que suponga que el desempleo es estacionario pero posiblemente no lineal y con el que obtendremos no sólo un mejor ajuste sino una mayor calidad predictiva que con un modelo lineal tradicional.

El resto de trabajo se estructura como sigue, en primer lugar, en la sección 2 se expone una justificación de la modelización no lineal de series temporales; en la sección 3 se concreta el ciclo modelizador en modelos no lineales, presentando los principales resultados relativos a las estrategias de modelización convenientemente desarrolladas en la literatura; en la sección 4 se aplica la metodología descrita a la serie de la tasa de desempleo de España. El artículo finaliza, sección 5, exponiendo las principales conclusiones que cabe extraer del ejercicio de economía aplicada efectuado.

## **2. JUSTIFICACIÓN DE LA MODELIZACIÓN NO LINEAL DE SERIES TEMPORALES**

La principal razón para tratar las series temporales como no lineales, nace de los diferentes regímenes o estados que pueden producirse en el comportamiento dinámico de las variables económicas a lo largo del tiempo. Para cada estado dinámico de la serie temporal se observará un comportamiento diferenciado de la media, varianza o autocorrelación.

Este comportamiento se refleja a menudo en la teoría económica. Algunos ejemplos de no linealidad comprenden a los procesos económicos con umbrales, las restricciones de capacidad en la producción, los desequilibrios persistentes debidos al racionamiento, las

restricciones institucionales como los tramos impositivos, los equilibrios múltiples y las asimetrías de diferente naturaleza, como las fluctuaciones cíclicas del empleo o el estudio de la evolución de la tasa de paro en función de los costes asimétricos de contratación y despido.

Si la no linealidad que presupone la teoría económica ha sido verificada con el estudio de los datos reales, se debe esperar que las ecuaciones estimadas, tengan también naturaleza no lineal. Sin embargo, la gran mayoría de modelos econométricos estimados son lineales, en muchos casos porque las ecuaciones pertinentes han sido remplazadas por una aproximación lineal que es considerada suficiente en la práctica. Otra razón para utilizar ecuaciones lineales es el deseo de evitar “teorías increíbles” -Sims (1980) -y llevar a cabo la modelización con los menores supuestos teóricos que sean posibles. Por último, el gran desarrollo de la teoría estadística para modelos lineales ha posibilitado un consistente avance de estrategias modelizadoras basadas en linealidad

En cambio, los modelos lineales no han estado ausentes de críticas. El usual rechazo de la constancia de los parámetros, indicativos de la llamada ruptura estructural ha sido corregido desechando el modelo o enmendándolo con variables dummies. Pero el comportamiento de una economía, un mercado, una empresa o una familia es demasiado complicado, como para poder ser adecuadamente descrito por una serie de ecuaciones lineales, por tanto uno debe esperar una ruptura del mismo a lo largo del tiempo. Las frecuentes rupturas que se producen en algunas series y la necesidad de aplicar varios grupos de variables de cambio hacen dudar de la credibilidad del modelo, siendo una respuesta más satisfactoria aplicar especificaciones que persiguen la modelización de los diferentes estados o regímenes que muestran las series a lo largo de su comportamiento dinámico.

No cabe duda que han sido los problemas que se han encontrado en los modelos lineales, además del espectacular crecimiento en el poder computacional disponible, la principal razón del aumento de la popularidad de la econometría no lineal, pero se ha visto acompañada por un incremento explosivo del número de estudios en los últimos veinte años, que han tratado desde una perspectiva no lineal los principales tópicos y dificultades de la econometría clásica.

La variedad de modelos de régimen cambiante es muy amplia. Una detallada revisión de los modelos existentes a lo largo de la bibliografía publicada hasta el momento puede consultarse en Casado (2003)<sup>1</sup>.

A pesar del gran número de modelos existentes podemos clasificarlos en dos grandes grupos. El primero estaría formado por aquellos que asumen que los distintos comportamientos o regímenes pueden ser determinados por una variable observable y por

---

<sup>1</sup> Sin ánimo de exhaustividad podemos citar entre los modelos considerados en la literatura los siguientes: *Threshold Autoregressive* (TAR), ver Tong (1978, 1990); *Self Exciting TAR* (SETAR), ver Tong y Lim (1980); *Smooth Transition Autoregressive* (STAR), ver Chang y Tong (1986), *Logistic STAR* (LSTAR), ver Teräsvirta y Anderson (1992), *Exponential STAR* (ESTAR), ver Teräsvirta (1994), *Multiple Regime STAR* (MRSTAR), ver van Dijk y Franses (1999), *Time-Varying STAR* (TVSTAR), ver Granger y Teräsvirta (1999); *Smooth Transition Error-Correction Model* (STECM), ver Granger y Swanson (1996); *Fractionally Integrated STAR* (FI-STAR), ver Van Dijk, Franses y Paap (2002).

tanto pueden ser concretados con absoluta certeza. Un segundo grupo de modelos son aquellos cuya hipótesis de partida es que los diferentes estados no son actualmente observables, pero están determinados por un proceso estocástico subyacente.

En las siguientes secciones, nos centraremos en la interpretación, especificación, test de linealidad, estimación, evaluación y predicción de aquellos modelos cuya variable de ruptura es observable, concentrando nuestro análisis en los que suponen una transmisión suave entre los distintos regímenes (STAR) y completando el ciclo modelizador para series temporales no lineales, recomendado por Teräsvirta (1994).

### 3. EL CICLO MODELIZADOR EN ANÁLISIS DE SERIES TEMPORALES NO LINEALES

Al igual que el procedimiento por etapas sugerido por Box y Jenkins (1970) para la modelización de series temporales lineales, Granger (1993) recomienda emplear el procedimiento “de lo específico a lo general” cuando consideramos la utilización de modelos de serie temporal no lineal, para describir las características de una variable determinada. Esto supone comenzar con un sencillo y restrictivo modelo, y proceder a la elaboración de uno más complejo sólo si los test y diagnósticos indican la falta de idoneidad del planteado. Los fundamentos del ciclo modelizador fueron inicialmente planteados por Teräsvirta (1994), consistiendo, ver la figura 1, en las siguientes etapas.

1. Especificación del modelo autorregresivo de orden  $p$  para la serie temporal objeto de estudio, usando un criterio apropiado de selección de modelos.
2. Elaboración de los test de linealidad.
3. Si la linealidad es rechazada, selección de la variable de transición  $s_t$  y de la forma funcional de la función de transición  $G(s_t; \gamma, c)$ .
4. Estimación de los parámetros del modelo elegido.
5. Evaluación del modelo, utilizando los mecanismos de validación específicos para modelos no lineales.
6. Modificación del modelo si fuera necesario.
7. Explotación del modelo con fines descriptivos. Predicción y análisis impulso-respuesta.

(Insertar Figura 1)

Las etapas 2 a la 7, serán discutidas con detalle en las siguientes secciones, pero antes realizaremos una breve introducción sobre las características y especificación de los modelos de transición suave más utilizados.

#### 3.1. Interpretación y especificación de los modelos STAR

Los modelos autorregresivos de transición suave (STAR, *Smooth Transition Autoregressive Model*) para una serie temporal univariante, observada desde  $t = 1 - p, 1 - (p - 1), \dots, -1, 0, 1, \dots, T - 1, T$ , se definen como:

$$y_t = (\phi_{1,0} + \phi_{1,1}y_{t-1} + \dots + \phi_{1,p}y_{t-p})(1 - G(s_t; \gamma, c)) + (\phi_{2,0} + \phi_{2,1}y_{t-1} + \dots + \phi_{2,p}y_{t-p})G(s_t; \gamma, c) + \varepsilon_t, \quad t = 1, 2, \dots, T \quad (1)$$

que expresamos, en forma reducida:

$$y_t = \phi_1' x_t (1 - G(s_t; \gamma, c)) + \phi_2' x_t G(s_t; \gamma, c) + \varepsilon_t \quad (2)$$

donde  $x_t = (1, \mathbf{x}_t')'$  con  $\mathbf{x}_t = (y_{t-1}, \dots, y_{t-p})'$  y  $\phi_i = (\phi_{i,0}, \phi_{i,1}, \dots, \phi_{i,p})'$ , siendo  $i = 1, 2$  para el caso de existencia de dos regímenes. Donde las perturbaciones aleatorias  $\varepsilon_t$  son independientes a la historia de la serie hasta el periodo  $t-1$ , que es expresado como  $E(\varepsilon_t / \Omega_{t-1}) = 0$ , donde  $\Omega_{t-1} = \{y_{t-1}, y_{t-2}, \dots, y_{1-(p-1)}, y_{1-p}\}$  y asumimos que la varianza condicional de  $\varepsilon_t$  es constante,  $E(\varepsilon_t^2 / \Omega_{t-1}) = \sigma^2$ .

La función de transición  $G(s_t; \gamma, c)$ , que puede adoptar diferentes formas funcionales, es una función continua, que toma valores en el intervalo (0,1), donde la variable de transición,  $s_t$  suele ser la variable endógena retardada,  $s_t = y_{t-d}$  para valores de  $d > 0$ . Si bien esta variable de transición puede también ser una variable exógena ( $s_t = z_t$ ), ver van Dijk y Franses (2000), o una tendencia lineal,  $s_t = t$  (ver Lin y Teräsvirta, 1994).

La interpretación de los modelos STAR puede ser analizada desde dos puntos de vista. Por una parte como modelos de dos regímenes cambiantes, cada uno asociado con un valor extremos de la función de transición  $G(s_t; \gamma, c) = 0$  y  $G(s_t; \gamma, c) = 1$ , donde la traslación de un modelo al otro es suave, o como un régimen continuo con diferentes valores de  $G(s_t; \gamma, c)$  que se mueven en el intervalo (0,1).

Son diferentes las formas funcionales que pueden adoptar la función de transición, pero nosotros centraremos nuestro estudios en dos. Una popular elección es la función logística de primer orden:

$$G(s_t; \gamma, c) = (1 + \exp\{-\gamma(s_t - c)\})^{-1}, \quad \gamma > 0 \quad (3)$$

En este caso el modelo STAR pasa a denominarse LSTAR (*Logistic STAR*). El parámetro  $c$  en (3) debe ser interpretado como el umbral o punto de inflexión entre los dos regímenes, ya que la función  $G(s_t; \gamma, c)$  es igual a 0.5 cuando  $s_t = c$ . El coeficiente  $\gamma$  mide la suavidad en la velocidad de cambio de un régimen a otro. Cuando  $\gamma$  toma valores elevados la velocidad de transición es elevada, y la función logística  $G(s_t; \gamma, c)$  se aproxima a la función indicador  $I[s_t > c]$ , definida como  $I[A] = 1$  si  $A$  es cierto y  $I[A] = 0$  en el resto de los casos y por tanto el cambio de  $G(s_t; \gamma, c)$  desde 0 hasta 1 se hace instantáneo en cuyo caso el modelo LSTAR se convierte en un modelo autorregresivo de

doble régimen (TAR) siempre que  $s_t = c$  y en un modelo *self-exciting TAR* (SETAR) cuando  $s_t = y_{t-d}$ . En el caso contrario, es decir, cuando  $\gamma$  tiende a cero la función logística hace que el modelo cambie de régimen de forma suave, siendo el caso límite aquel en el que  $\gamma = 0$  (la función  $G(s_t; \gamma, c) = 0.5$ ) y el modelo LSTAR se reduce a un AR con parámetros  $\phi_j = \frac{(\phi_{1,j} + \phi_{2,j})}{2}$ ,  $j = 0, 1, \dots, p$ .

En algunos casos es más apropiado especificar la función de transición, de tal forma que los regímenes estén asociado con pequeños o grandes valores de  $s_t$ , de nuevo en relación a  $c$ , de forma distinta a como lo hace el LSTAR, esto se logra con una función de translación exponencial (ESTAR), que se define como,

$$G(s_t; \gamma, c) = 1 - \exp\left\{-\gamma(s_t - c)^2\right\} \quad \gamma > 0 \quad (4)$$

### 3.2. Tests de linealidad en modelos STAR

Antes de construir un modelo no lineal es aconsejable averiguar si realmente un modelo lineal sería adecuado para caracterizar las relaciones económicas objeto de análisis. Pero como demostraremos más adelante, los test de linealidad también son útiles a la hora de especificar el tipo de modelo no lineal, así como en la fase de validación, ya que es recomendable someter a los residuos del modelo especificado a dichos test, para verificar que las características no lineales de la serie original han sido capturadas adecuadamente.

La hipótesis nula de linealidad puede ser expresada como la igualdad de los parámetros autorregresivos de los dos regímenes del modelo STAR, expresión (2), esto es:

$$\begin{aligned} H_0 : \phi_1 &= \phi_2 \\ H_1 : \phi_1 &\neq \phi_2 \end{aligned} \quad (5)$$

Pero la realización de este test se complica por la presencia de parámetros no especificados bajo la hipótesis nula. La hipótesis nula no restringe los valores de los parámetros de la función de transición,  $\gamma$  y  $c$ , cuando la hipótesis nula es cierta, y la verosimilitud se ve afectada por la presencia de los valores de ambas medidas. Este problema de no identificación en los test de linealidad puede ser reformulado de forma más clara a partir del siguiente contraste

$$\begin{aligned} H'_0 : \gamma &= 0 \\ H'_1 : \gamma &\neq 0 \end{aligned} \quad (6)$$

que también da lugar a un modelo lineal.

Este inconveniente fue tratado por Luukkonen et al (1988a), cuya propuesta para su solución fue sustituir la función  $G(s_t; \gamma, c)$  por una aproximación de Taylor. En la ecuación reparametrizada, el problema de identificación no está presente y la linealidad

puede ser comprobada con un estadístico de multiplicadores de Lagrange que sigue asintóticamente una distribución  $\chi^2$  bajo la hipótesis nula.

Partiendo del modelo LSTAR:

$$y_t = \phi_1' x_t + (\phi_2 - \phi_1)' x_t G(s_t; \gamma, c) + \varepsilon_t \quad (7)$$

y asumiendo que  $\varepsilon_t \sim N.I.D(0, \sigma^2)$ , Luukkonen et al. (1988a) aconseja aproximar la función logística  $G(s_t; \gamma, c) = (1 + \exp\{-\gamma(s_t - c)\})^{-1}$ ,  $\gamma > 0$ , con una aproximación de Taylor de primer orden en torno a  $\gamma = 0$ , dando como resultado la regresión auxiliar:

$$y_t = \beta_0' x_t + \beta_1' x_t s_t + e_t \quad (8)$$

donde  $\beta_i = (\beta_{i,0}, \beta_{i,1}, \dots, \beta_{i,p})'$ ,  $i = 0, 1$ ;  $e_t = \varepsilon_t + (\phi_2 - \phi_1)' x_t R_1(s_t; \gamma, c)$ , siendo  $R_1(s_t; \gamma, c)$  el término resto del desarrollo de Taylor. Bajo la hipótesis nula  $R_1(s_t; \gamma, c) \equiv 0$  y  $e_t = \varepsilon_t$  y el resto no afecta a las propiedades de las perturbaciones aleatoria. Los parámetros  $\beta_i$ ,  $i = 0, 1$  en la regresión (8) están desempeñando la función de los parámetros del modelo (7), tal que  $\gamma = 0$  implica  $\beta_{0,j} \neq 0$  y  $\beta_{1,j} \neq 0$  para  $j = 0, \dots, p$  pudiendo expresar, por tanto, a partir de (8) la hipótesis (5) y (6) en (7) como<sup>2</sup>:

$$\begin{aligned} H_0'' : \beta_1 &= 0 \\ H_1'' : \beta_1 &\neq 0 \end{aligned} \quad (9)$$

El test estadístico para contrastar esta restricción, denominado  $LM_1$ , sigue una distribución asintótica  $\chi^2$  con  $p+1$  grados de libertad bajo la hipótesis nula de linealidad.

Luukkonen et al. (1988b) demuestra que el test  $LM_1$  no tiene potencia suficiente en los casos en los que solo varía el término constante, entre los diferentes regímenes del modelo STAR, es decir:

$$\begin{aligned} \phi_{1,0} &\neq \phi_{2,0} \quad \text{y} \\ \phi_{1,j} &= \phi_{2,j} \quad \text{para } j = 1, \dots, p \end{aligned} \quad (10)$$

desarrollando para su solución el test  $LM_3$  con una aproximación de  $G(s_t; \gamma, c)$  generada con un desarrollo de Taylor de tercer orden cuya regresión auxiliar queda especificada como:

$$y_t = \beta_0' x_t + \beta_1' x_t s_t + \beta_2' x_t s_t^2 + \beta_3' x_t s_t^3 + e_t \quad (11)$$

donde  $e_t = \varepsilon_t + (\phi_2 - \phi_1)' x_t R_3(s_t; \gamma, c)$  y de nuevo  $\beta_{0,0}$  y  $\beta_i, i = 1, 2, 3$  actúan como los parámetros  $\phi_{1,\phi}, \phi_{2,\phi}, \gamma$  y  $c$  pudiendo expresar ahora la hipótesis  $H_0' : \gamma = 0$  como:

---

<sup>2</sup> El orden  $p$ , número de retardos introducidos en los modelos LM de contraste de linealidad, son el resultado de la especificación, estimación y validación de un modelo linear autorregresivo (AR) previo usando un apropiado criterio de selección de modelos.



$$\begin{aligned} H_0'' : \beta_1 = \beta_2 = \beta_3 = 0 \\ H_1'' : \text{al menos un } \beta_i \neq 0 \end{aligned} \quad (12)$$

Bajo la hipótesis nula, el test LM sigue una distribución asintótica  $\chi^2$  con  $3(p+1)$  grados de libertad.

La expresión de  $\beta_i, i=1,2,3$ , en (11) en término de  $\phi_1, \phi_2, \gamma$  y  $c$  es tal que sólo los parámetros  $\beta_{1,0}, \beta_{2,0}$  y  $\beta_{3,0}$  representan a los interceptos  $\phi_{1,0}$  y  $\phi_{2,0}$ . Una versión más parsimoniosa del estimador  $LM_3$  puede ser obtenida aumentando la estimación auxiliar (8) con regresores como  $s_t^2$  y  $s_t^3$ , tal que:

$$y_t = \beta_0' x_t + \beta_1' x_t s_t + \beta_{2,0} s_t^2 + \beta_{3,0} s_t^3 + e_t \quad (13)$$

y se efectúa el test:

$$\begin{aligned} H_0'' : \beta_1 = 0 \\ \beta_{2,0} = \beta_{3,0} = 0 \\ H_1'' : \text{al menos un } \beta_i \neq 0 \end{aligned} \quad (14)$$

El test estadístico resultante se denomina  $LM_3^e$  el cual sigue asintóticamente una distribución  $\chi^2$  con  $p+3$  grados de libertad bajo la hipótesis nula de linealidad.

Los test de linealidad para modelos ESTAR fueron expuesto inicialmente por Saikkonen y Luukkonen (1988) sugiriendo chequear la regresión

$$y_t = \beta_0' x_t + \beta_1' x_t s_t + \beta_2' x_t s_t^2 + e_t \quad (15)$$

donde  $e_t = \varepsilon_t + (\phi_2 - \phi_1)' x_t R_2(s_t; \gamma, c)$ .

Esta ecuación se obtiene a partir del desarrollo de Taylor de primer orden para la expresión (2) utilizando como expresión de la variable de transición (4), donde la restricción  $\gamma = 0$ , es analizada con la hipótesis,

$$\begin{aligned} H_0'' : \beta_1 = \beta_2 = 0 \\ H_1'' : \text{al menos un } \beta_i \neq 0 \end{aligned} \quad (16)$$

El estadístico  $LM_2$ , sigue asintóticamente, bajo la hipótesis nula, una distribución  $\chi^2$  con  $2(p+1)$  grados de libertad.

Por último, Escribano y Jordá (1999) demuestran que la aproximación de Taylor de primer orden para la función exponencial no es suficiente para capturar los dos puntos de inflexión de esta función y sugirieron un desarrollo de Taylor de segundo orden,

$$y_t = \beta_0' x_t + \beta_1' x_t s_t + \beta_2' x_t s_t^2 + \beta_3' x_t s_t^3 + \beta_4' x_t s_t^4 + e_t \quad (17)$$

Al igual que en los casos anteriores en éste lo que se contrasta es la hipótesis

$$\begin{aligned} H_0'' : \beta_1 = \beta_2 = \beta_3 = \beta_4 = 0 \\ H_1'' : \text{al menos un } \beta_i \neq 0 \end{aligned} \quad (18)$$

En este último caso, el test  $LM_4$  sigue una distribución asintótica  $\chi^2$  con  $4(p+1)$  grados de libertad bajo la hipótesis nula.

Antes de finalizar este breve repaso de los test de linealidad en modelos STAR debemos hacer referencia al hecho de que estos test, en el caso de que en la serie existan problemas de heteroscedasticidad o sea numerosa la presencia de outliers, no son eficientes. En dichas circunstancias, por tanto, es preciso plantear una especificación robusta de los test LM.

Wooldridge (1990) ha diseñado la especificación de los test que deben ser utilizados en presencia de heteroscedasticidad, sin ser necesario especificar la forma funcional que sigue dicha heteroscedasticidad. Un ejemplo para el test  $LM_3$  sería desarrollar las siguientes etapas:

- Estimar  $y_t$  en  $x_t$  y obtener los residuos  $\hat{\varepsilon}_t$ .
- Estimar la regresión auxiliar  $x_t s_t^i, i = 1, 2, 3$ , sobre  $x_t$  y calcular los residuos  $\hat{r}_t$ .
- Calcular la regresión de 1 frente en  $\hat{\varepsilon}_t \hat{r}_t$ . El estadístico LM es el resultado de la suma explicada de esta regresión.

En el caso de la presencia de outliers, al igual que para la heteroscedasticidad planteamos, a modo de ejemplo, el procedimiento de elaboración del estimador robusto para el estadístico  $LM_3$ , a partir del resultado de  $nR^2$  donde el coeficiente de determinación  $R^2$  es el derivado de la regresión de los residuos ponderados  $\hat{w}_r(\hat{r}_t) \hat{r}_t$  sobre los regresores ponderados  $\hat{w}_x(x_t) \otimes (x_t', x_t' s_t, x_t' s_t^2, x_t' s_t^3)'$ , siendo  $\otimes$  el producto elemento por elemento. La ponderación de los residuos  $\hat{w}_r(\hat{r}_t)$  y de los regresores  $\hat{w}_x(x_t)$ , la obtenemos de un modelo  $AR(p)$  bajo la hipótesis nula. El estadístico LM resultante, sigue asintóticamente una distribución  $\chi^2$  con  $3(p+1)$  grados de libertad.

### 3.3. Estimación

Una vez que la linealidad ha sido rechazada, es necesario como etapa previa del proceso de estimación, seleccionar la variable de transición  $s_t$ , así como la forma funcional más adecuada de la función de transición  $G(s_t; \gamma, c)$  (en nuestro caso LSTAR vs ESTAR).

Una manera tradicional de elegir la variable de transición ha sido a partir del p-value del estadístico  $LM_3$ , ya que al estar integrada la especificación ESTAR de (15) en (11), puede elegirse la variable  $s_t$  de forma independiente a la forma funcional de  $G(s_t; \gamma, c)$ . Este procedimiento sugiere que la apropiada variable de transición del modelo STAR puede ser determinada computando el  $LM_3$  estadístico para las variables de transición candidatas eligiendo aquella en la que el p-value es más pequeño. La racionalidad de este procedimiento, confirmada en Teräsvirta (1994), se fundamenta en el hecho de que los test deberían tener más potencia en los casos en los que el modelo alternativo está correctamente especificado, y esto es, cuando hemos utilizado la variable de transición adecuada.

Una vez que la linealidad ha sido rechazada y la variable de transición ha sido elegida, pero antes de proceder a la estimación, debemos seleccionar la forma funcional de la función de transición  $G(s_i; \gamma, c)$ . Son dos los test que tradicionalmente se ha utilizado con este fin, por una parte el propuesto por Teräsvirta (1994), fundamentado en el contraste de una secuencia de hipótesis nulas anidadas,  $H_{03} : \beta_3 = 0$ ,  $H_{02} : \beta_2 = 0/\beta_3 = 0$  y  $H_{01} : \beta_1 = 0/\beta_3 = \beta_2 = 0$  sobre la regresión (11) eligiendo un modelo ESTAR si el  $p$ -value de  $H_{02}$  es inferior y un LSTAR en el resto de los caso. Por otra parte Escribano y Jordá (1999) propusieron un test basado en la comprobación de dos hipótesis separadas efectuadas a partir de la expresión (17):

$$\begin{aligned} H_{0E} : \beta_2 = \beta_4 = 0 \\ H_{0L} : \beta_1 = \beta_3 = 0 \end{aligned} \quad (19)$$

y comparando los  $p$ -values de ambos estimadores. El LSTAR será el seleccionado si el mínimo  $p$ -value es obtenido para  $H_{0L}$ .

Una vez seleccionada la variable y función de transición la siguiente fase en el ciclo modelizador descrito en la figura 1 es la estimación de los parámetros del modelo STAR mediante un procedimiento de mínimos cuadrados no lineales (NLS)<sup>3</sup>, donde lo que se persigue es la obtención de los parámetros  $\theta = (\phi'_1, \phi'_2, \gamma, c)'$  mediante la minimización de,

$$\hat{\theta} = \arg \min_{\theta} Q_T(\theta) = \arg \min_{\theta} \sum_{t=1}^T (y_t - F(x_t; \theta))^2 \quad (20)$$

donde  $F(x_t; \theta)$  es el esqueleto del modelo no lineal, es decir,

$$F(x_t; \theta) = \phi'_1 x_t (1 - G(s_t; \gamma, c)) + \phi'_2 x_t G(s_t; \gamma, c). \quad (21)$$

Wooldridge (1994) y Pötscher y Prucha (1997) demostraron que mínimos cuadrados no lineales genera estimadores consistentes y asintóticamente normales,

$$\sqrt{T}(\hat{\theta} - \theta_0) \rightarrow N(0, C) \quad (22)$$

donde  $\theta_0$  es el verdadero valor del parámetro.

La matriz de covarianza asintótica  $C$  de  $\hat{\theta}$  puede ser estimada consistentemente por el producto de  $\hat{A}_T^{-1} \hat{B}_T \hat{A}_T^{-1}$ , donde  $\hat{A}_T$  es el Hessiano de  $\hat{\theta}$ , esto es:

$$\hat{A}_T = -\frac{1}{T} \sum_{t=1}^T \nabla^2 q_t(\hat{\theta}) = \frac{1}{T} \sum_{t=1}^T (\nabla F(x_t; \hat{\theta}) \nabla F(x_t; \hat{\theta})' - \nabla^2 F(x_t; \hat{\theta}) \hat{\varepsilon}) \quad (23)$$

con  $q_t(\hat{\theta}) = (y_t - F(x_t; \hat{\theta}))^2$  y  $\hat{B}_T$ :

$$\hat{B}_T = \frac{1}{T} \sum_{t=1}^T \nabla q_t(\hat{\theta}) \nabla q_t(\hat{\theta})' = \frac{1}{T} \sum_{t=1}^T \hat{\varepsilon}_t^2 \nabla F(x_t; \hat{\theta}) \nabla F(x_t; \hat{\theta})' \quad (24)$$

<sup>3</sup> Recordamos que bajo la hipótesis de que las perturbaciones aleatorias están normalmente distribuidas, mínimos cuadrados no lineales (NLS) es equivalente a máxima verosimilitud y por tanto NLS puede ser interpretado como una estimación cuasi máximo verosímil.

La estimación de los parámetros puede ser realizada por diferentes tipos de algoritmos de convergencia y procedimientos de optimización, por ejemplo los desarrollados en Hamilton (1994) o Hendry (1995).

### 3.4. Diagnostic Checking (EVALUATION)

Antes de que el modelo STAR sea admitido como válido, debe ser sometido a una exhaustiva evaluación, la cual debe incluir además de las clásicas etapas de los modelos lineales otros test específicos para modelos no lineales como son los de no autocorrelación, no-linealidad restante o constancia de los parámetros.

#### (a) Contraste de no autocorrelación.

A partir del modelo STAR considerado en (2) y de la forma funcional del esqueleto del modelo (21) un test LM de orden  $q$ th para la dependencia serial de  $\varepsilon_t$  puede ser obtenido como  $nR^2$ , donde  $R^2$  es el coeficiente de determinación de la regresión de  $\hat{\varepsilon}_t$  en

$$\nabla F(x_t; \hat{\theta}) = \partial F(x_t; \hat{\theta}) / \partial \theta \quad (25)$$

Siendo  $\theta = (\phi'_1, \phi'_2, \gamma, c)'$  y  $q$  los retardos de los residuos  $\hat{\varepsilon}_{t-1}, \dots, \hat{\varepsilon}_{t-q}$ , donde el símbolo circunflejo indican que han sido estimados bajo la hipótesis nula de independencia serial de  $\varepsilon_t$ . El resultado del estadístico, se distribuye asintóticamente como una  $\chi^2$  con  $q$  grados de libertad.

#### (b) Contraste de no-linealidad restante

Hasta el momento siempre hemos considerado que la no-linealidad supone la existencia de dos regímenes. En cambio, la configuración general de los modelos STAR permite la existencia de varios modelos de régimen múltiple que contienen sub-modelos con distintas velocidades de ajustes y puntos de ruptura.

Los modelos STAR que incorporan más de un régimen deben ser clasificados en dos grupos, por una parte aquellos en los que cada estado está caracterizado por una única variable de transición  $s_t$ , y por otra, aquellos, en los que una combinación de varias variables  $s_{1t}, \dots, s_{mt}$  determinan el régimen en el que nos encontramos. En el primer caso debemos de describir (2) para obtener un modelo de tres regímenes como sigue:

$$y_t = \phi'_1 x_t + (\phi_2 - \phi_1)' x_t G_1(s_t; \gamma_1, c_1) + (\phi_3 - \phi_2)' x_t G_2(s_t; \gamma_2, c_2) + \varepsilon_t \quad (26)$$

Si asumimos que  $c_1 < c_2$ , los parámetros autorregresivos en este modelo cambian lentamente desde  $\phi_1$  a través de  $\phi_2$  a  $\phi_3$  ante incrementos del valor de  $s_t$ , cuando la primera de las funciones  $G_1$  cambia desde 0 hasta 1, seguida de cambios similares de  $G_2$ . De forma más general se puede definir un conjunto de  $m-1$  parámetros de transición suave,  $\gamma_1, \dots, \gamma_{m-1}$  y un conjunto de  $m-1$  parámetros de localización  $c_1, \dots, c_{m-1}$  para lograr un modelo STAR con  $m$  regímenes de cambio,

$$y_t = \phi'_1 x_t + (\phi_2 - \phi_1)' x_t G_1(s_t) + \dots + (\phi_m - \phi_{m-1})' x_t G_{m-1}(s_t) + \varepsilon_t \quad (27)$$

donde  $G_j(s_t) \equiv G_j(s_t; \gamma_j, c_j)$ ,  $j = 1, \dots, m-1$  puede adoptar la forma de la función logística (3) transformándose (6) en un modelo SETAR de régimen múltiple cuando el parámetro de velocidad de transición  $\gamma$ , toma valores elevados.

La formalización del modelo STAR para el caso en el que los regímenes están determinados por una combinación de diferentes variables puede realizarse por “encapsulamiento” de dos modelos STAR de doble régimen de la siguiente manera:

$$y_t = \{\phi'_1 x_t (1 - G_1(s_{1t}; \gamma_1, c_1))\} + \{\phi'_2 x_t (G_1(s_{1t}; \gamma_1, c_1))\} [1 - G_2(s_{2t}; \gamma_2, c_2)] + \{\phi'_{31} x_t (1 - G_1(s_{1t}; \gamma_1, c_1))\} + \{\phi'_4 x_t (G_1(s_{1t}; \gamma_1, c_1))\} [G_2(s_{2t}; \gamma_2, c_2)] + \varepsilon_t \quad (28)$$

donde la relación entre  $y_t$  y sus retardos viene dada por una combinación lineal de cuatro modelos AR, cada uno asociado con una particular combinación de  $G_1(s_{1t})$  y  $G_2(s_{2t})$  que toman valores entre 0 y 1.

El MRSTAR, ver van Dijk y Franses (1999) tiene en cuenta como máximo un modelo con cuatro regímenes diferentes, pero aplicando el principio de encapsulamiento de forma repetida, el modelo puede ser extendido a  $2^m$  regímenes.

Eitrheim y Teräsvirta (1996) desarrollaron un estadístico LM para chequear la existencia de un LSTAR de doble régimen frente al especificado en (26), donde la hipótesis nula de existencia de un modelo de doble régimen es igual a

$$H'_0 : \gamma_2 = 0 \quad \text{ó} \quad H'_0 : \phi_3 = \phi_2 \quad (29)$$

Al igual que en los test de linealidad planteados, estos test presentan problemas de parámetros no identificados, cuestión que es resuelta de nuevo, especificando un desarrollo de Taylor siguiendo una aproximación de tercer orden para remplazar la función de transición  $G_2(s_t; \gamma_2, c_2)$  por su aproximación, que debe ser expresada como:

$$y_t = \beta'_0 x_t + (\phi_2 - \phi_1)' x_t G_1(s_t; \gamma_1, c_1) + \beta'_1 x_t s_t + \beta'_1 x_t s_t^2 + \beta'_1 x_t s_t^3 + e_t \quad (30)$$

donde  $\beta_i, i = 0, 1, 2, 3$ , son funciones de los parámetros  $\phi_1, \phi_2, \phi_3, \gamma_2$  y  $c$ . De forma que la hipótesis (29) en (26) se puede identificar, considerando la expresión (30), con:

$$H''_0 : \beta_1 = \beta_2 = \beta_3 = 0 \quad (31)$$

Al menos un  $\beta_i \neq 0$

El test estadístico resultante tradicionalmente definido como  $LM_{AMR,3}$  sigue una distribución asintótica  $\chi^2$  con  $3(p+1)$  grados de libertad.

### (C) Test de constancia de los parámetros

La no linealidad es una de las muchas características que las series temporales pueden poseer, en cambio, otra de las más comunes peculiaridades es la existencia de cambios estructurales. Durante tiempo ambas, se han visto como dos formas distintas de analizar una misma realidad, sin embargo, en los últimos años se han diseñado una nueva tipología

de modelos no lineales, denominados modelos STAR de variación temporal (TVSTAR) que incluyen ambas características.

Los modelos TVSTAR, suponen que  $y_t$  siguen un modelo STAR todo el tiempo, pero con un cambio suave en los parámetros autorregresivos en ambos regímenes, es decir de  $\phi_1 \leftrightarrow \phi_2$ , a  $\phi_3 \leftrightarrow \phi_4$ , para  $G(s_{1t}; \gamma_1, c_1) = 0$  y  $G(s_{1t}; \gamma_1, c_1) = 1$ , respectivamente.

Su expresión general es la siguiente:

$$y_t = \phi_1(t)'x_t(1 - G_1(s_{1t}; \gamma_1, c_1)) + \phi_2(t)'x_t(G_1(s_{1t}; \gamma_1, c_1)) + \varepsilon_t \quad (32)$$

siendo:

$$\begin{aligned} \phi_1(t) &= \phi_1[1 - G_2(t; \gamma_2, c_2)] + \phi_3 G_2(t; \gamma_2, c_2) \\ \phi_2(t) &= \phi_2[1 - G_2(t; \gamma_2, c_2)] + \phi_4 G_2(t; \gamma_2, c_2) \end{aligned}$$

Por tanto, contrastando  $\gamma_2 = 0$  en (32) contrastamos la constancia de los parámetros en un modelo STAR de doble régimen. El test LM correspondiente, obtenido a partir de una aproximación de Taylor de tercer orden de  $G_2(t; \gamma_2, c_2)$ , se denomina  $LM_{C,3}$  y sigue una distribución asintótica  $\chi^2$  con  $6(p+1)$  grados de libertad.

### 3.5. Predicción

Los modelos no lineales pueden ser utilizados con varios propósitos. A veces el principal objetivo es solamente obtener una adecuada descripción del comportamiento dinámico que presenta una determinada variable. Pero a menudo, un objetivo adicional es utilizar el modelo para predecir valores futuros de la serie temporal, siendo explotada esa predicción para evaluar la calidad del modelo STAR, comparándolo con un modelo lineal tradicional.

Consideremos el caso donde  $y_t$  es explicado por un modelo STAR con  $s_t = y_{t-1}$ :

$$y_t = F(x_t; \theta) + \varepsilon_t \quad (33)$$

donde el esqueleto del modelo está formado por:

$$F(x_t; \theta) = \phi_1'x_t(1 - G(y_{t-1}; \gamma, c)) + \phi_2'x_t G(y_{t-1}; \gamma, c) \quad (34)$$

siendo  $x_t = (1, y_{t-1}, \dots, y_{t-p})'$ .

La predicción puntual de  $y_{t+h}$  para el periodo  $t$  viene dada por su media condicional, debiéndose expresar para un horizonte de predicción  $h$  como:

$$\hat{y}_{t+h/t} = E[y_{t+h}/\Omega_t] \quad (35)$$

donde  $\Omega_t$  representa la historia de la serie temporal hasta el momento, incluyendo la observación del periodo  $t$ .

A partir de (35), y teniendo en cuenta que  $E[e_{t+1}/\Omega_t] = 0$ , la predicción con un horizonte igual a 1, puede ser obtenida como:

$$\hat{y}_{t+1/t} = E[y_{t+1}/\Omega_t] = F(y_t; \theta) \quad (36)$$

En cambio cuando el horizonte de predicción es superior a 1, la predicción puntual se complica. Supongamos el caso en el que  $h = 2$ :

$$\hat{y}_{t+2/t} = E[y_{t+2}/\Omega_t] = E[F(y_{t+1}; \theta)/\Omega_t] \quad (37)$$

En este caso, la esperanza condicional  $E(\bullet)$  no puede ser intercambiada con el operador no lineal  $F(\bullet)$ , es decir:

$$E[F(\bullet)] \neq F(E[\bullet]) \quad (38)$$

Por tanto, el valor esperado de la función no-lineal no es igual a la función calculada del valor esperado, esto es:  $E[F(y_{t+1}; \theta)/\Omega_t] \neq F(E[y_{t+1}/\Omega_t]; \theta)$ .

Cuando el horizonte de predicción es superior a un periodo, la expresión de  $\hat{y}_{t+h/t}$ , no es posible, ya que la expresión  $E[y_{t+h}/\Omega_t]$  requiere una dimensión  $h-1$ . Aunque en un primer momento se plantearon técnicas de integración numérica para solucionar este problemas en los últimos años han sido los métodos de Monte Carlo y los iterativos bootstrap los que se han afianzado para aproximar la esperanza condicional de  $E[y_{t+h}/\Omega_t]$ , calculando la predicción de Monte Carlo para un horizonte de predicción de orden 2 como,

$$\hat{y}_{t+2/t}^{mc} = \frac{1}{k} \sum_{i=1}^k F(\hat{x}_{t+2/t}^{(i)}; \theta) \quad (39)$$

donde  $k$  es algún valor elevado y los valores de  $\varepsilon_{t+1}$  en  $(\hat{x}_{t+2/t}^{(i)})$  son obtenidos de la supuesta distribución de  $\varepsilon_t$ .

La predicción obtenida por el método iterativo es muy similar y la única diferencia es que  $\varepsilon_{t+1}$  se obtiene por sustitución de los residuos del modelo estimado  $\hat{\varepsilon}_t, t=1, \dots, T$ . teniendo la ventaja de que no es necesario suponer ningún tipo de distribución para  $\varepsilon_t$ . Otra de las ventajas de los métodos Monte Carlo e iterativo es que  $F(\hat{x}_{t+2/t}^{(i)}; \theta)$  constituye la forma de la función de densidad de predicción, permitiendo ser utilizada par construir intervalos de confianza.

Hasta el momento hemos asumido que los parámetros del modelo STAR son conocidos, pero en la práctica son estimados y esto añade un componente de incertidumbre adicional a la estimación que puede ser recogida añadiendo a las predicciones Monte Carlo o Iterativas un componente promedio sobre los diferentes valores de los parámetros. Estos es, obteniendo:

$$\hat{y}_{t+2/t}^{mc} = \frac{1}{kr} \sum_{i=1}^k \sum_{j=1}^r F(\hat{x}_{t+2/t}^{(i)}; \theta^{(j)}) \quad (40)$$

donde  $\theta^{(j)}$  se obtiene de la distribución de los parámetros estimados  $\hat{\theta}$ .

### 3.6. Funciones de impulso respuesta

Las funciones de impulso respuesta utilizadas en modelos lineales, son definidas como la diferencia entre dos realizaciones de  $y_{t+h}$  definidas para idéntica historia de la serie temporal,  $t-1$ , que denotamos como,  $\omega_{t-1}$ . En la primera, el proceso es sometido a un shock de tamaño  $\delta$  en el momento  $t$ , mientras que la segunda no es “shockeada” por ningún impulso. Todos los shocks en los periodos intermedios entre  $t$  y  $t+h$ , son iguales a cero en ambas realizaciones, por tanto podemos representar las funciones de impulso respuesta tradicionales ( $TI$ ) como:

$$TI_y(h, \delta, \omega_{t-1}) = E[y_{t+h} | \varepsilon_t = \delta, \varepsilon_{t+1} = \dots = \varepsilon_{t+h} = 0, \omega_{t-1}] - E[y_{t+h} | \varepsilon_t = 0, \varepsilon_{t+1} = \dots = \varepsilon_{t+h} = 0, \omega_{t-1}] \quad (41)$$

para  $h = 1, 2, 3, \dots$

Simetría e independencia histórica son dos propiedades de las funciones  $TI$  en series temporales lineales. La primera supone que un shock de tamaño,  $-\delta$  tiene exactamente el efecto opuesto a un shock de tamaño  $+\delta$ , y la respuesta es proporcional al tamaño del shock. La segunda elimina la dependencia del impacto en relación a la historia particular  $\omega_{t-1}$ . Ambas propiedades son fácilmente observables en el  $TI$  de un proceso AR(1).

$$TI_y(h, \delta, \omega_{t-1}) = \phi_1^h \delta, \quad h = 0, 1, 2, \dots \quad (42)$$

En cambio si hacemos extensible los  $TI$  a modelos no lineales observamos como algunas de estas propiedades no se mantienen. Así, suponiendo un modelo no lineal simple (SETAR):

$$y_t = \begin{cases} \phi_{1,1} y_{t-1} + \varepsilon_t, & \text{si } y_{t-1} \leq 0 \\ \phi_{1,2} y_{t-1} + \varepsilon_t, & \text{si } y_{t-1} > 0 \end{cases} \quad (43)$$

se obtiene:

$$TI_y(1, \delta, \omega_{t-1}) = \begin{cases} \phi_{1,1} \delta & \text{si } y_{t-1} + \delta \leq 0, \quad y_{t-1} \leq 0 \\ \phi_{1,1} \delta + \delta_{1,2} (\phi_{1,1} - \phi_{1,2}) y_{t-1} & \text{si } y_{t-1} + \delta \leq 0, \quad y_{t-1} > 0 \\ \phi_{1,2} \delta + \delta_{1,1} (\phi_{1,2} - \phi_{1,1}) y_{t-1} & \text{si } y_{t-1} + \delta > 0, \quad y_{t-1} \leq 0 \\ \phi_{1,2} \delta & \text{si } y_{t-1} + \delta > 0, \quad y_{t-1} > 0 \end{cases} \quad (44)$$

A partir de (44) puede deducirse claramente que en modelos no lineales, la respuesta ante un shock depende del signo y del tamaño del mismo, así como de la historia del proceso. Por tanto ni la simetría, ni la dependencia histórica van a ser propiedades de los modelos no lineales, al tener valores distintos para  $y_{t-1} + \delta > 0$  y  $y_{t-1} - \delta \leq 0$ .

Al ser las funciones de impulso respuesta una buena herramienta para considera el comportamiento dinámico de los modelos no lineales estimados y dadas las peculiaridades de los mismos, Koop et al (1996) diseñaron una útil generalización del concepto de funciones de respuesta impulso para modelo no lineales definida como  $GI$ , que se define, para un shock específico  $\varepsilon_t = \delta$  y una historia  $\omega_{t-1}$ , como:



$$GI_y(h, \delta, \omega_{t-1}) = E[y_{t+h} | \varepsilon_t = \delta, \omega_{t-1}] - E[y_{t+h} | \omega_{t-1}] \quad (45)$$

para  $h = 0, 1, 2, \dots$

En el  $GI$ , la esperanza de  $y_{t+h}$  dada por un shock  $\delta$  ocurrida en  $t$  está condicionada por la historia y el shock, siendo una función de  $\delta$  y  $\omega_{t-1}$  que puede ser entendida como diferentes consecuciones de la variable aleatoria  $\varepsilon_t$  y  $\Omega_{t-1}$ . Koop et al (1996) puntualiza que  $GI_y(h, \delta, \omega_{t-1})$  es en si mismo una sucesión de variables aleatorias, definidas como:

$$GI_y(h, \delta, \Omega_{t-1}) = E[y_{t+h} | \varepsilon_t, \Omega_{t-1}] - E[y_{t+h} | \Omega_{t-1}] \quad (46)$$

A partir de esta interpretación general del  $GI$  como una variable aleatoria, varias versiones condicionadas pueden ser consideradas de interés. Se puede tener en cuenta sólo un historia particular  $\omega_{t-1}$  considerando el  $GI$  como una variable aleatoria en términos de  $\varepsilon_t$ :

$$GI_y(h, \varepsilon_t, \omega_{t-1}) = E[y_{t+h} | \varepsilon_t, \omega_{t-1}] - E[y_{t+h} | \omega_{t-1}] \quad (47)$$

o, por el contrario, determinar para las diferentes historias el  $GI$  para un shock determinado a priori  $\varepsilon_t = \delta$ , considerándolo como una variable aleatoria en términos de la historia  $\Omega_{t-1}$ .

En la práctica, se suele calcular el  $GI$  condicionada a una particular combinación de (A y B) shocks e historias respectivamente  $GI_y(h, A, B)$ . Por ejemplo para todas las historias en un régimen particular y con shocks negativos.

Por último Koop et al (1996) nos demuestra, que en el caso de los modelos STAR, la expresión analítica para la esperanza condicional de  $GI$  cuando  $h > 1$ , no esta disponible, y por tanto debemos utilizar simulación estocástica para obtener el valor de las diferentes respuestas como veremos en el desarrollo empírico.

Quizás sea la medición de la persistencia, la mayor de las capacidades explicativas de las funciones de impulso respuesta. Recordemos que decimos que un shock  $\varepsilon_t = \delta$  es transitorio para una historia  $\omega_{t-1}$  si  $GI_y(h, \delta, \omega_{t-1})$  tiende a 0 cuando  $h$  tiende a infinito, en caso contrario, decimos que el shock es persistente. Potter (1995) y Koop et al (1996) sugieren que la dispersión de la distribución  $GI_y(h, \varepsilon_t, \Omega_{t-1})$  en un horizonte finito puede ser interpretada como una medida de la persistencia de los shocks. Esto nos permitirá comparar y medir densidades (persistencia) de  $GI$ s condicionados a tipos de shocks (negativos-positivos) teniendo en cuenta a su vez las diferencias generadas por los regímenes del modelo STAR especificado.

## 4. MODELIZACIÓN DEL DESEMPLEO ESPAÑOL. UN ENFOQUE NO LINEAL

El estudio del desempleo español debe tener en cuenta tres características de la serie, que pueden observarse fácilmente en la figura 2, la cual representa la tasa de desempleo registrado por el INEM (instituto nacional de empleo) comprendida en el periodo enero de 1982 hasta diciembre del 2002.

(Insertar figura 2)

- Naturaleza asimétrica con respecto al ciclo económico. Crecimiento acelerado en las etapas recesivas y un decrecimiento lento y moderado en las épocas de expansión, característica que debe ser captada con modelos que contemplen regímenes diferenciados y comportamientos contracíclicos.

- Persistencia. Aunque tradicionalmente relacionada, siguiendo el concepto de histéresis, con la no estacionariedad, y comprobada con los test de raíces unitarias, Skalin y Teräsvirta (2002) comprueban aplicando el test de Dickey-Fuller a diferentes series simuladas no lineales que los modelos de cambio de régimen no proporcionan evidencia adicional frente a la hipótesis de raíz unitaria comparada con el correspondiente modelo lineal. Por tanto la búsqueda de raíces unitarias por un lado y la no-linealidad (asimetría) por otro, utilizando los mismos datos, no debe ser entendida necesariamente como una contradicción. Por tanto nosotros en este análisis seguimos el enfoque de Bianchi y Zoega (1998) y Skalin y Teräsvirta (2002) suponiendo que la tasa de desempleo es globalmente estacionaria pero posiblemente no-lineal y localmente no estacionaria.

- En cambio, la especificación de los modelos STAR en (1) no contempla otra de las características observables en la figura 2, el pronunciado comportamiento estacional, (observable en los gráficos de las figuras 3 y 4) mostrando valores mayores en otoño e invierno (noviembre-febrero) que a finales de los meses de verano (julio-octubre). Aunque algunos trabajos asumen que el componente estacional cambia a lo largo del tiempo, el estudio de la estacionalidad de la serie en nuestro caso, nos demuestra que no esta sujeto a cambios (figura 4) durante los periodos expansivo/recesivos y por tanto debe ser recogido en el modelo no lineal utilizando variables dummies mensuales,  $D_{s,t}$ ,  $s = 1, \dots, 11$ , donde  $D_{s,t} = 1$  si la observación  $t$  corresponde al mes  $s$  y  $D_{s,t} = 0$  para el resto de los casos, que no están influenciadas por los ciclos expansivos/recesivos.<sup>4</sup>

(Insertar figuras 3 y 4)

Utilizando la muestra desde enero de 1985 para analizar la serie descrita, y reservando los últimos tres años para predicción, comenzamos el ciclo modelizador, enunciado en la sección anterior, especificando un modelo lineal autorregresivo. Llevando a cabo un comportamiento de lo específico a lo general, observamos como tanto el AIC como el BIC, indicaban que un modelo AR(4) era el más apropiado. En cambio su comportamiento

---

<sup>4</sup> Se ha comprobado aplicando el contraste de Hylleberg et al (1990) la naturaleza determinista del componente estacional del desempleo español, quedando demostrada la idoneidad de las variables ficticias como instrumento para recoger el componente estacional. (Franses, 1996).

parsimonioso y la existencia de autocorrelación en los residuos nos ha llevado a aumentar los retardos a 13, eliminando aquellas variables cuyos parámetros no eran significativos. De esta manera, hemos obtenido:

$$\begin{aligned} \Delta y_t = & -0.122_{(0.061)} - 0.002_{(0.003)} y_{t-1} + 0.337_{(0.071)} \Delta y_{t-1} + 0.213_{(0.069)} \Delta y_{t-2} + 0.168_{(0.068)} \Delta y_{t-3} - 0.163_{(0.059)} \Delta y_{t-5} + \\ & + 0.153_{(0.060)} \Delta y_{t-8} + 0.237_{(0.065)} \Delta y_{t-12} - 0.093_{(0.067)} \Delta y_{t-13} + 0.284_{(0.050)} D_{1,t} + 0.168_{(0.049)} D_{2,t} + \\ & + 0.144_{(0.056)} D_{3,t} - 0.047_{(0.044)} D_{4,t} - 0.048_{(0.044)} D_{5,t} + 0.167_{(0.048)} D_{6,t} + 0.292_{(0.049)} D_{8,t} + \\ & + 0.401_{(0.054)} D_{9,t} + 0.321_{(0.056)} D_{10,t} + 0.191_{(0.044)} D_{11,t} + \hat{\varepsilon}_t \end{aligned}$$

$$\hat{\sigma}_e = 0.1091, SK = 0.1241(0.2478)_p, EK = 3.5300(0.0727)_p, JB = 2.5830(0.2745)_p,$$

$$LM_p(1) = 0.9845, LM_p(4) = 0.6616, LM_p(8) = 0.7036, LM_p(12) = 0.2736,$$

$$ARCH_p(1) = 0.2484, ARCH_p(2) = 0.4730, AIC = -4.220, BIC = -3.884$$

$p$  = Valor de la probabilidad

Los valores entre paréntesis son las desviaciones típicas obtenidas por mínimos cuadrados ordinarios de los parámetros estimados, y  $\hat{\varepsilon}_t$  los residuos de la regresión en el momento  $t$ ,  $\hat{\sigma}_e$  la desviación típica de los residuos,  $SK$  la asimetría,  $EK$  el nivel de curtosis,  $JB$  el test de normalidad Lomnichi-Jarque-Bera de los residuos,  $ARCH(q)$  el test  $LM$  de no existencia de heteroscedasticidad autorregresiva condicional de orden  $q$  y  $LM(p)$  el test de Breusch-Godfrey de no existencia de autocorrelación de orden  $p$ .

El modelo lineal muestra un adecuado comportamiento de los residuos, ya que no están correlacionados, y no existe heteroscedasticidad autorregresiva condicional en ellos. No se detecta asimetría significativa, pero sí una leve curtosis que no llega a ser relevante a la hora de aceptar la normalidad con el test  $JB$ .

La siguiente etapa consiste en chequear la linealidad frente al STAR utilizando los estadísticos  $LM$  explicados en el apartado 3.2 (ver tabla 1). Como nuestro interés se centra en el comportamiento del desempleo con respecto al ciclo económico, debemos utilizar una variable de transición del modelo STAR, que refleje las propiedades de los periodos de expansión y recesión. Por tanto el cambio mensual de la tasa de desempleo, es una variable inapropiada como proxy del ciclo económico ya que refleja las fluctuaciones estacionales en la tasa de desempleo. Siguiendo a Skalin y Teräsvirta (2002), utilizaremos la diferencia duodécima como variable de transición  $s_t = \Delta_{12} y_{t-d} \equiv y_{t-d} - y_{t-d-12}$  seleccionando como valor máximo de  $d$  seis.

(Insertar Tabla 1)

En primer lugar centrándonos, en los test estándares, observamos como para el contraste  $LM_1$  y  $LM_3^e$ , cualquiera de los niveles puede ser seleccionada como variable de transición al 5% no pudiendo decir lo mismo en el caso de los test  $LM_3$  y  $LM_4$ , para los que

parecería más adecuado utilizar cualquiera de los tres primeros. En cambio el resultado general de los test, parece indicar un claro comportamiento no lineal de la serie, no viéndose influida esta decisión por la presencia de outliers, ya que los test robustos de no linealidad nos llevaría a adoptar la misma conclusión.

En una tercera etapa, y una vez rechazada la linealidad, seleccionamos la variable de transición  $s_t$  y la forma funcional de transición  $G(s_t; \gamma, c)$ . Es la diferencia primera estacional,  $\Delta_{12}y_{t-1}$ , la que muestra un menor valor de la probabilidad en el estadístico  $LM_3$ . En cuanto a la especificación óptima y basándonos en la regla de decisión de Teräsvirta (1994) las tres variantes del test sugieren un modelo LSTAR como el más apropiado considerando cualquiera de las variable de transición, confirmando los estadísticos de Escribano-Jordá la decisión adoptada.

(Insertar Tabla 2)

Por tanto hasta el momento y con la combinación de las tablas 1 y 2, la variable  $\Delta_{12}y_{t-1}$  ha sido elegida como variable de transición de un modelo LSTAR. En cambio y con la finalidad de seleccionar la más adecuada hemos estimado diferentes modelos LSTAR para  $s_t = \Delta_{12}y_{t-d}$ , siendo  $d = 1, 2, 3$ , obteniendo en las diferentes etapas de evaluación y predicción los mejores resultados para  $d=1$  y por tanto confirmando la variable seleccionada por los procedimientos tradicionales.

Partiendo de un modelo autorregresivo de orden 13 para cada uno de los regímenes, realizamos el proceso de estimación del LSTAR, obteniendo:

$$\begin{aligned} \Delta y_t = & \underset{(0.056)}{-0.006} + \underset{(0.038)}{0.202} D_{1,t} - \underset{(0.039)}{0.050} D_{3,t} - \underset{(0.039)}{0.231} D_{4,t} - \underset{(0.042)}{0.234} D_{5,t} - \underset{(0.043)}{0.157} D_{7,t} + \\ & + \underset{(0.045)}{0.160} D_{8,t} + \underset{(0.050)}{0.381} D_{9,t} + \underset{(0.050)}{0.288} D_{10,t} + \underset{(0.048)}{0.190} D_{11,t} + \left[ \underset{(0.004)}{-0.007} y_{t-1} + \underset{(0.098)}{0.238} \Delta y_{t-2} + \right. \\ & + \underset{(0.099)}{0.110} \Delta y_{t-4} + \underset{(0.096)}{0.149} \Delta y_{t-8} - \underset{(0.084)}{0.167} \Delta y_{t-9} - \underset{(0.093)}{0.101} \Delta y_{t-10} + \underset{(0.077)}{0.165} \Delta y_{t-12} \Big] \cdot \\ & \cdot [1 - G(\Delta_{12}y_{t-1}; \hat{\gamma}, \hat{c})] + \left[ \underset{(0.003)}{0.002} y_{t-1} + \underset{(0.073)}{0.372} \Delta y_{t-1} + \underset{(0.078)}{0.132} \Delta y_{t-2} + \underset{(0.082)}{0.163} \Delta y_{t-3} - \right. \\ & \left. - \underset{(0.071)}{0.105} \Delta y_{t-8} - \underset{(0.070)}{0.089} \Delta y_{t-12} - \underset{(0.073)}{0.098} \Delta y_{t-12} \right] G(\Delta_{12}y_{t-1}; \hat{\gamma}, \hat{c}) + \hat{\varepsilon}_t \end{aligned}$$

$$G(\Delta_{12}y_{t-1}; \hat{\gamma}, \hat{c}) = (1 + \exp\{4.743(\Delta_{12}y_{t-1} - 0.668)/\hat{\sigma}_{\Delta_{12}y_{t-1}}\})^{-1}$$

$$\hat{\sigma}_e = 0.1034, \hat{\sigma}(LSTAR/AR) = 0.94, SK = 0.049(0.3929)_p, EK = 3.393(0.1404)_p,$$

$$JB = 1.236(0.5390)_p, LM_p(1) = 0.8376, LM_p(4) = 0.4041, LM_p(8) = 0.4255,$$

$$LM_p(12) = 0.2144, ARCH_p(1) = 0.3542, ARCH_p(2) = 0.3435, AIC = -4.251,$$

$$BIC = -3.79$$

$p$  = Valor de la probabilidad

Donde  $\hat{\sigma}(LSTAR/AR)$  representa el ratio de la desviación típica de los residuos del modelo LSTAR frente al autorregresivo, disminuyendo un 6% la desviación del LSTAR con respecto al AR. Además esta disminución es suficiente para compensar el incremento del número de parámetros (de 20 a 26) ya que el LSTAR es preferido al AR con respecto al AIC, aunque no podamos decir lo mismo con respecto al BIC. La disminución de la asimetría y del exceso de curtosis, nos lleva a aceptar con mayor intensidad la normalidad de los residuos. Los test LM y ARCH, no rechazan la hipótesis nula para cualquiera de los retardos analizados.

En las figuras 5 y 6 se representan, respectivamente, los residuos del modelo LSTAR correspondiente a la serie objeto de estudio y la función de distribución kernel de los residuos del modelo LSTAR estimado.

(Insertar Figuras 5 y 6)

Terminamos, la etapa de evaluación del LSTAR con los test de autocorrelación, no linealidad restante y constancia de los parámetros (ver tabla 3), obteniendo que el modelo seleccionado se encuentra validado.

(Insertar Tabla 3)

La función de transición  $G(\Delta_{12}y_{t-1}; \hat{\gamma}, \hat{c})$  estimada frente a la variable de transición  $\Delta_{12}y_{t-1}$  se muestra en la figura 7. La estimación de  $\hat{\gamma}$  y  $\hat{c}$ , muestra que la transición de la función logística desde 0 a 1 tiene lugar para valores de  $\Delta_{12}y_{t-1}$  comprendidos entre -2 y 0.5, siendo la velocidad de transición moderada e igual a 4.74 y el punto de inflexión de la transición  $\Delta_{12}y_{t-1} = -0.668$ .

(Insertar Figura 7)

Los tres últimos años de la serie temporal, desde Enero de 2000 hasta Diciembre de 2002, fueron reservados para evaluar la calidad predictiva del modelo AR y LSTAR estimado. Para cada momento del tiempo desde Enero de 2000 y hasta Enero de 2002 computamos la predicción de 1 a 12 horizontes tanto para el modelo AR como para el LSTAR. Estas últimas han sido calculadas mediante el método iterativo bootstrap especificado en el apartado 3.5.

La predicción se ha realizado de una forma estática, sin actualizar los parámetros estimados cuando una nueva observación estaba disponible. La tabla 4 contiene la evaluación de la predicción haciendo uso de algunos de los criterios más generalizados. A partir del criterio MPE, error de predicción media, el modelo LSTAR, a pesar de ser pesimista con respecto a la disminución del desempleo, por ser su valor de predicción mayor que el real, presenta una desviación menor que el autorregresivo en un horizonte temporal corto (1 y 2 períodos). En cambio el modelo AR es más constante en la calidad de la predicción y esta menos influido por la dimensión temporal en la que es medida. Comparando el MSPE, error de predicción cuadrático medio, de ambos modelos obtenemos una conclusión muy similar, ya que aunque la calidad del modelo LSTAR es

mejor en el corto plazo es más recomendable la utilización del lineal para horizontes temporales de largo plazo.

(Insertar Tabla 4)

Para conocer las propiedades dinámicas del modelo STAR estimado, analizamos el comportamiento del modelo ante diferentes shocks, mediante la elaboración de funciones generalizadas de impulso respuesta especificadas en el apartado 3.6. Calculamos los GIs, como fueron definidos en (45), teniendo en cuenta el tamaño muestral del modelo estimado para diferentes shocks normalizados  $\delta/\hat{\sigma}_\varepsilon = \pm 3, \pm 2.9, \dots, \pm 0.2, \pm 0.1$ , donde  $\hat{\sigma}_\varepsilon$  representa la desviación típica de los residuos estimados en el modelo LSTAR especificado. Para cada combinación de historia (180) y shocks (60) calculamos un  $GI_{\Delta y}(h, \delta, \omega_{t-1})$  con un horizonte de  $h = 0, 1, \dots, 60$ . La esperanza condicional de (45) es estimada como la media de 1000 realizaciones de  $\Delta y_{t+h}$ , obtenidas a través de iteraciones del modelo LSTAR especificado, con y sin la utilizations del shock seleccionado y utilizando aleatoriamente muestras de los residuos estimados en el modelo no lineal especificado. Los impulsos respuesta para la variable desempleo en niveles es obtenido por agregación de los impulsos respuestas en primeras diferencias, esto es:

$$GI_y(h, \delta, \omega_{t-1}) = \sum_{i=0}^h GI_{\Delta y}(i, \delta, \omega_{t-1}).$$

Los diferentes GIs son calculados para A shocks y B historias,  $GI_y(h, A, B)$  respectivamente, donde A representa el conjunto de todos los positivos y negativos shocks, y B las historias para las que la función de transición  $G(\Delta_{12}y_{t-1}; \gamma, c)$  estimada es mayor (recesión) o menor (expansión) que 0.5.

Por último las densidades son calculadas con un estimador Kernel Nadaraya-Watson, utilizando  $\phi(\delta/\hat{\sigma}_\varepsilon)$  como ponderación de  $GI_y(h, \delta, \omega_{t-1})$  donde  $\phi(z)$  representa la distribución típica de probabilidad normal, y con los HDR (Highest Density Regions) utilizando la metodología tipificada en Hansen (1996).

(Insertar Figura 9)

La figura 9 muestra los HDRs para las distribuciones de  $GI_y(h, A, B)$  con horizontes  $h=0, 3, 6, \dots, 60$ , pudiendo observar cómo los shocks que ocurren durante las recesiones tienden a alcanzar su máxima respuesta en el sexto mes. En cambio en los periodos de expansión los shock no lograrán el máximo valor hasta pasado un año. Por otra parte el efecto de shocks negativos es en general inferior al de los positivos, lo que nos llevarían a demostrar el escaso poder de las políticas de creación de empleo y la rigidez del mercado de trabajo, independientemente del momento de ciclo económico en el que nos encontremos. Además resaltar, la menor asimetría de los impactos en periodos de recesión con respecto a los de expansión.

Por último, se puede contemplar comparando las gráficas, una de las razones fundamentales que justifican este trabajo, y es que si se realiza políticas de creación de

empleo en periodos de recesión, éstas tiene un efecto mucho menor, que si en épocas de expansión ocurre algún shock que perjudica a su evolución, quedando demostrada la naturaleza contracíclica del desempleo.

## 5. CONCLUSIONES

La tasa de desempleo español y en general el mercado de trabajo es, como señalábamos en la introducción, una variable que debe ser analizada teniendo en cuenta sus peculiaridades. Su comportamiento asimétrico, incrementándose más rápidamente en las etapas de recesión que lo que disminuye en las de expansión genera, como ha sido comprobado en este trabajo, dos comportamientos totalmente diferenciados de la serie, recomendándose la utilización de modelos no lineales para su especificación

La demostración en este trabajo de que el desempleo español tiene un comportamiento totalmente distinto en las fases expansivas y recesivas, requiere de los organismos que toman decisiones, políticas activas y pasivas de desempleo diferenciadas en función de la fase del ciclo económico en el que nos encontremos. El diseño de políticas de empleo a largo plazo e independientes de la situación coyuntural en la que se encuentre la economía, no sólo pueden ser ineficaces sino que contribuyen a perjudicar al comportamiento esperado del mercado de trabajo. Tenemos que tener en cuenta que no nos encontramos ante un mismo patrón en el que sólo varía su evolución positiva/negativa sino que como se observa en el modelo estimado se trata de dos modelos autorregresivos totalmente diferenciados.

La correcta especificación no lineal de la serie, no ha sido solamente confirmada, por los test de linealidad, sino que tanto en la evaluación del modelo estimado como en la predicción a corto plazo, se ha demostrado una sustancial mejora en todos los contrastes realizados con respecto al modelo lineal, destacando la idoneidad a la hora de seleccionar dos regímenes y no ordenes superiores en la evolución de la tasa de desempleo. Permanencia estructural (no siendo necesaria la estimación por procedimientos TVSTAR), inexistencia de autocorrelación o ausencia de heteroscedasticidad autorregresiva condicional, son algunas de las óptimas cualidades que el modelo no-lineal presenta.

Analizando el comportamiento dinámico del modelo STAR estimado con las funciones de impulso respuesta, se confirma la naturaleza contracíclica y asimétrica del desempleo español, quedando demostrada además la escasa capacidad de las políticas de creación de empleo debido a la rigidez del mercado laboral español.

Por último destacar que los resultados obtenidos para el desempleo español no difieren de los alcanzados en otros países, aunque si hay que resaltar, siguiendo el artículo de van Dijk, Teräsvirta y Franses (2002) la mayor magnitud de las respuestas que las políticas de empleo tienen en Estados Unidos, quedando demostrada la rigidez relativa del mercado laboral español y el menor margen de maniobra de las políticas de empleo en España.

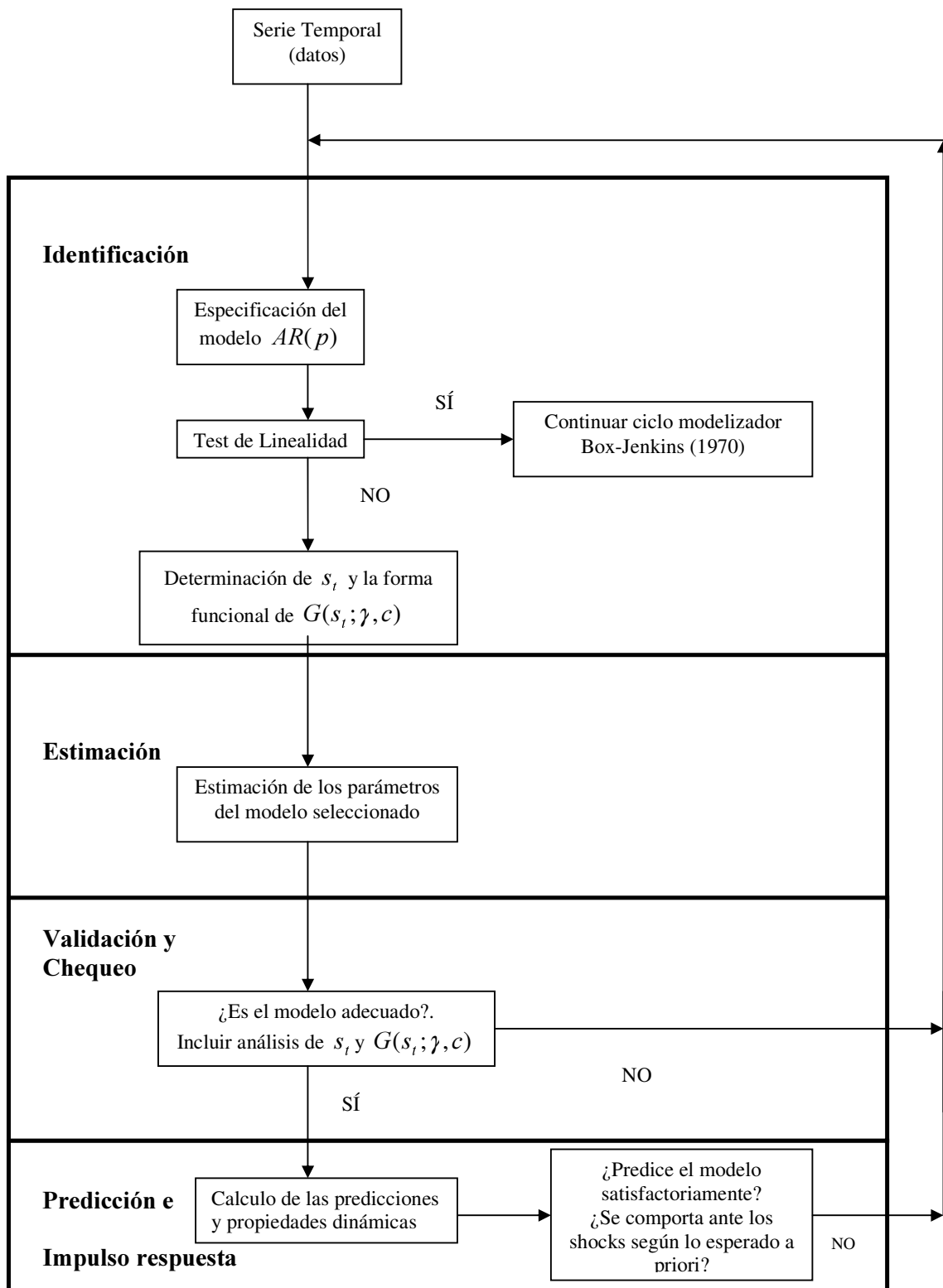
## REFERENCIAS BIBLIOGRÁFICAS

- Bentolila, S., and Bertola, G. (1990), "Firing costs and labour demand: How bad is eurosclerosis?" *Review of Economics and Statistics*. 57, 381-402.
- Bianchi, M., and Zoega, G. (1998), "Unemployment Persistence: Does the Size of the Shock Matter?" *Journal of Applied Econometrics*. 13, 283-304.
- Blanchard, O.J., Rhee, C. and Summers, L. (1990), "The Stock Market, Profit and Investment." MIT; U Rochester; Harvard U.
- Blanchard, O. J., and Summers, L.H. (1987), "Hysteresis in unemployment." *European Economic Review* 31(1987): 288-295.
- Box, G.E.P. and Jenkins, G.M. (1970), *Time Series Analysis: Forecasting and Control*. San Francisco: Holden Day.
- Brannas, K. and Ohlsson, H. (1999), "Asymmetric Time Series and Temporal Aggregation." *Review of Economics and Statistics*, 2, 341-44.
- Caner, M. and Hansen, B.E. (2001), "Threshold Autoregression with a Unit Root." *Econometrica*, 69, 1555-1596.
- Casado, J.M. (2003), "Análisis coyuntural de la industria española. Asimetría, persistencia y no linealidad de la tasa de desempleo industrial". *Documento de trabajo*. Universidad de Zaragoza.
- Chang y Tong (1986), "On estimating thresholds in autoregressive models". *Journal of Time Series Analysis* 7, 179-90.
- Eitrheim, O. and Teräsvirta, T. (1996), "Testing the Adequacy of Smooth Transition Autoregressive Models." *Journal of Econometrics*, 74, 59-75.
- Escribano, A. and Jordá, O. (1999), "Improved Testing and Specification of Smooth Transition Regression Models." En *Nonlinear Time Series Analysis of Economic and Financial Data*, Dordrecht: Kluwer Academic Press, 289-319.
- Franses, P.H. (1996), *Periodicity and Stochastic Trends in Economic Time Series*, Oxford: Oxford University Press.
- Granger, C. W. J. (1993), "Strategies for Modelling Nonlinear Time-Series Relationships." *Economic Record*, 69, 233-238.
- Granger, C.W.J. and Swanson, N.R. (1996), "Future developments in the study of cointegrated variables". *Oxford Bulletin of Economics and Statistics* 58, 537-53.
- Granger C.W.J. and Teräsvirta, T. (1999), "A simple nonlinear time series model with misleading linear properties." *Economics Letters* 62, 161-5
- Hamilton, J. D. (1994), *Time Series Analysis*. Princeton University Press, 1994.

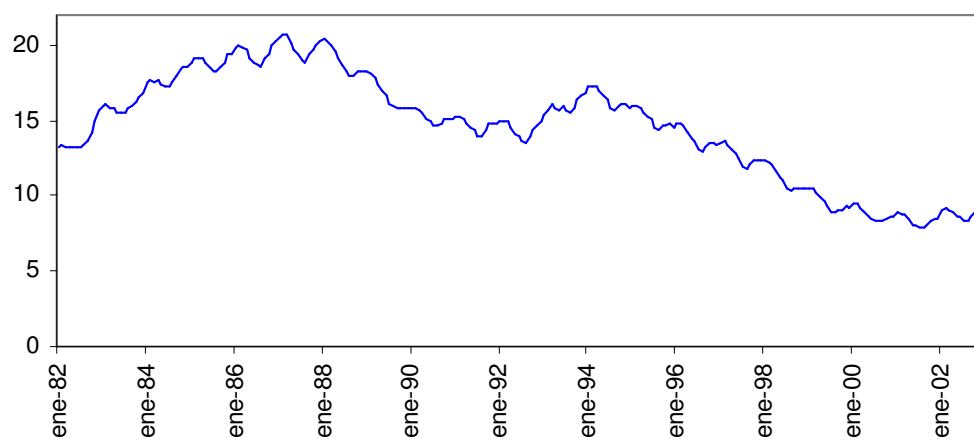


- Hansen, B.E. (1996), "Inference when a Nuisance Parameter is not Identified under the Null Hypothesis". *Econometrica*, 64,413-430.
- Hansen, B. E. (1997), "Inference in TAR Models." *Studies in Nonlinear Dynamics and Econometrics*, 2, 1-14.
- Hendry, D. F. (1995), *Dynamic econometrics*. Oxford: Oxford University Press.
- Hylleberg, S., Engle, R.F., Granger, C.W.J. and Yoo, B.S. (1990), Seasonal Integration and Cointegration, *Journal of Econometrics*, 49, 215-238.
- Koop, G., and Potter, S.M. (1999), "Dynamic Asymmetries in U.S. Unemployment" *Journal of Business and Economic Statistics*, 17, 298-312.
- Koop, G., Pesaran, M.H. and Potter, S.M. (1996), "Impulse Response Analysis in Nonlinear Multivariate Models". *Journal of Econometrics*, 74, 119-147.
- Lin, C. F. J. and Teräsvirta, T. (1994), "Testing the Constancy of Regression Parameters against Continuous Structural Change." *Journal of Econometrics*, 62, 211-228.
- Lindbeck, A. and Snower, D. (1985), *Explanations of Unemployment*, Oxford: Oxford University Press.
- Lindbeck, A. and Snower, D.J. (1988), "Cooperation, Harassment, and Involuntary Unemployment: An Insider-Outsider Approach." *American Economic Review*, 78, 167-188.
- Lindgren, G. (1978), "Markov Regime Models for Mixed Distributions and Switching Regressions." *Scandinavian Journal of Statistics*, 5, 81-91.
- Luukkonen, R., Saikkonen, P. and Teräsvirta, T. (1988a), "Testing Linearity against Smooth Transition Autoregressive Models." *Biometrika*, 75, 491-499.
- Luukkonen, R., Saikkonen, P. and Teräsvirta, T. (1988b), "Testing Linearity in Univariate Time Series Models." *Scandinavian Journal of Statistics*, 16, 161-175.
- Montgomery, A. L., Zarnovitz, V., Tsay, R.S. and Tiao, G.C. (1998), "Forecasting the U.S. Unemployment Rate." *Journal of the American Statistical Association*, 93, 478-493.
- Neftçi, S. (1984), "Are Economic Time Series Asymmetric over the Business Cycle?" *Journal of Political Economy*, 92, 307-328.
- Parker, R. E. and Rothman, P. (1997), "The Current Depth-of-Recession and Unemployment Rate Forecats". *Studies in Nonlinear Dynamics and Econometrics*, 2, 151-158.
- Potter, S.M. (1995), "Nonlinear Models of Economic Fluctuations. In *Macroeconometrics Developments, Tensions and Prospects*. Boston: Kluwer, 517-560.
- Potscher, B. M. and Prucha. I.R. (1997), *Dynamic Nonlinear Econometric Models: Asymptotic Theory*. Berlin: Springer-Verlag.

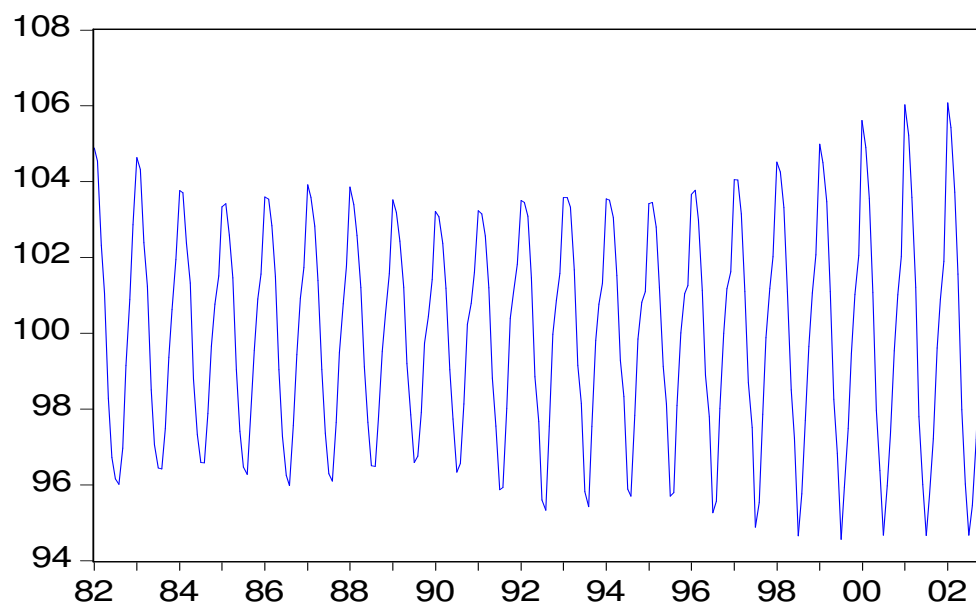
- Rothman, P. (1998), "Forecasting Asymmetric Unemployment Rates". *Review of Economics and Statistics*, 80, 164-168.
- Sachs, J. D. (1986), "High Unemployment in Europe: Diagnosis and Policy Implications." Department of Economics. Harvard University.
- Saikkonen, P. and Luukkonen, R. (1988), "Lagrange Multiplier Tests for Testing Non-linearities in Time Series Models." *Scandinavian Journal of Statistics*, 15, 55-68.
- Skalin, J. and Teräsvirta, T. (2002), "Modeling Asymmetries and Moving Equilibria in Unemployment Rates". *Macroeconomic Dynamics*, 6, 202-241.
- Sims, C.A. (1980), "Macroeconomics and Reality". *Econometrica*, 40, 1-48.
- Teräsvirta, T. (1994), "Specification, Estimation, and Evaluation of Smooth Transition Autoregressive Models". *Journal of the American Statistical Association*, 89, 208-218.
- Teräsvirta, T. and Anderson, H.M. (1992), "Characterizing Nonlinearities in Business Cycles Using Smooth Transition Autoregressive Models". *Journal of Applied Econometrics*. Suppl 7, 119-136.
- Tong, H., (1978), "On a threshold model". *Pattern Recognition and Signal Processing*, Amsterdam: Sijthoff & Noordhoff, 101-41
- Tong, H. (1990), *Non-Linear Time Series: A Dynamical Systems Approach*. Oxford: Oxford University Press.
- Tong, H. and Lim, K.S. (1980) "Threshold autoregressions, limit cycle, and data". *Journal of the Royal Statistical Society B* 42, 245-92.
- van Dijk, D. and Franses, P.H. (1999), "Modeling Multiple Regimes in the Business Cycle." *Macroeconomic Dynamics*, 3, 311-340.
- van Dijk, D. and Franses, P.H. (2000), *Nonlinear time series models in empirical finance*. Cambridge: Cambridge University Press.
- van Dijk, D., Teräsvirta, T. and Franses, P.H. (2002), "Smooth Transition Autoregressive Models. A Survey of Recent Developments." *Econometric Reviews*, 21, 1-47.
- van Dijk, D., Franses, P.H. and Paap, R. (2002), "A Nonlinear Long Memory Model with an Application to US Unemployment." *Journal of Econometrics*, 110, 135-165.
- Wooldridge, J. M. (1990), "A Unified Approach to Robust, Regression-Based Specification Tests". *Econometric Theory*, 6, 17-43.
- Wooldridge, J. M. (1994), Estimation and Inference for Dependent Processes. In *Handbook of Econometrics*. Vol. IV. Elsevier Science, 2639-2738.



**Figura 1:** El ciclo modelizador en series temporales no lineales

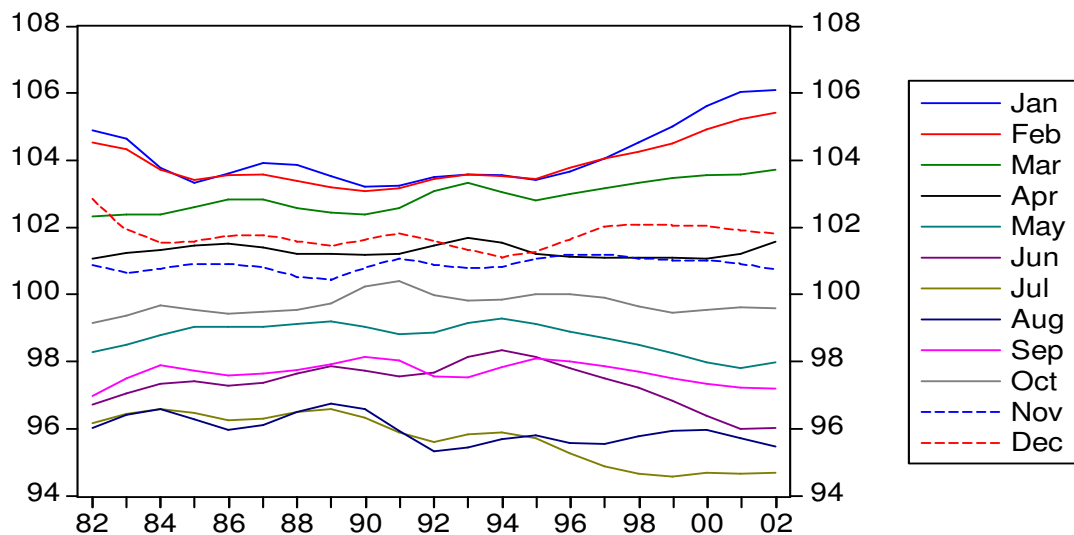


**Figura 2:** Evolución de la tasa de desempleo mensual en España. Período: Enero 1982 a Diciembre 2002.



**Figura 3:** Componente estacional de la tasa de desempleo obtenido aplicando TRAMO/SEATS

**Figura 4:** Componente estacional de la tasa de desempleo por periodos.



**Tabla 1:** Test tipo LM para no linealidad STAR

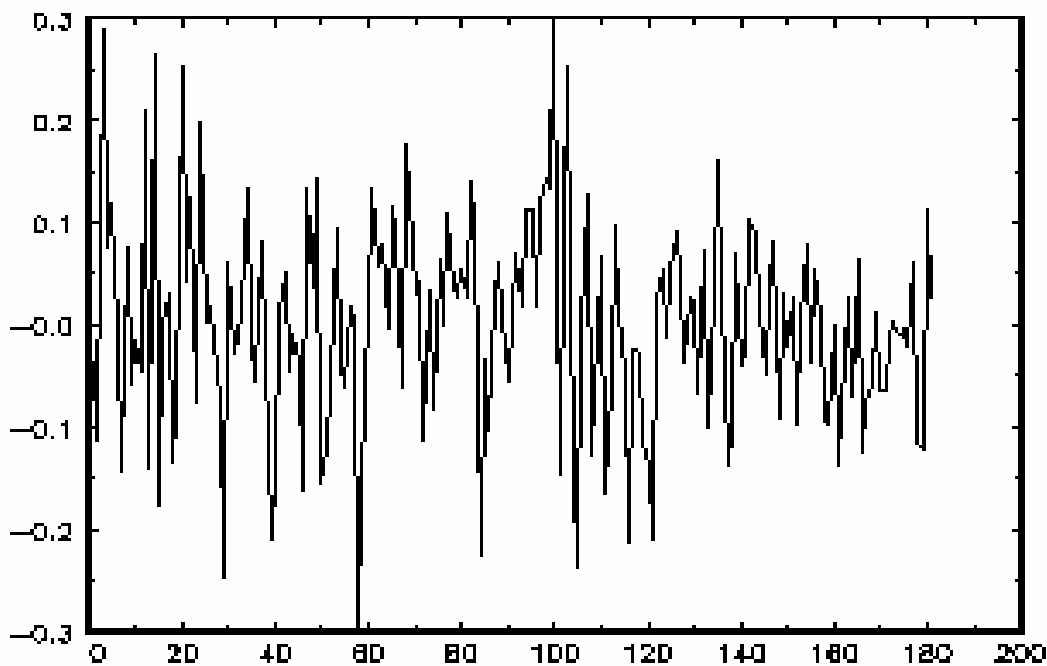
Variable Transición ( $s_t$ )	Test Estándar				Test Robusto Outliers			
	$LM_1$	$LM_3$	$LM_3^e$	$LM_4$	$LM_1$	$LM_3$	$LM_3^e$	$LM_4$
$\Delta_{12}y_{t-1}$	0.000	0.001	0.000	0.023	0.007	0.007	0.004	0.089
$\Delta_{12}y_{t-2}$	0.000	0.006	0.000	0.019	0.009	0.013	0.009	0.050
$\Delta_{12}y_{t-3}$	0.001	0.028	0.001	0.022	0.011	0.039	0.013	0.152
$\Delta_{12}y_{t-4}$	0.004	0.312	0.006	0.128	0.020	0.253	0.026	0.341
$\Delta_{12}y_{t-5}$	0.007	0.378	0.015	0.082	0.036	0.393	0.060	0.114
$\Delta_{12}y_{t-6}$	0.012	0.285	0.020	0.364	0.043	0.189	0.059	0.333

Nota: *p-values* de los test tipo LM para la no-linealidad de la tasa de desempleo española, Enero 1985-Diciembre 1999. Los test son aplicados sobre un modelo AR(13) para la primera diferencia, incluyendo la variable en niveles retardada y las ficticias estacionales.

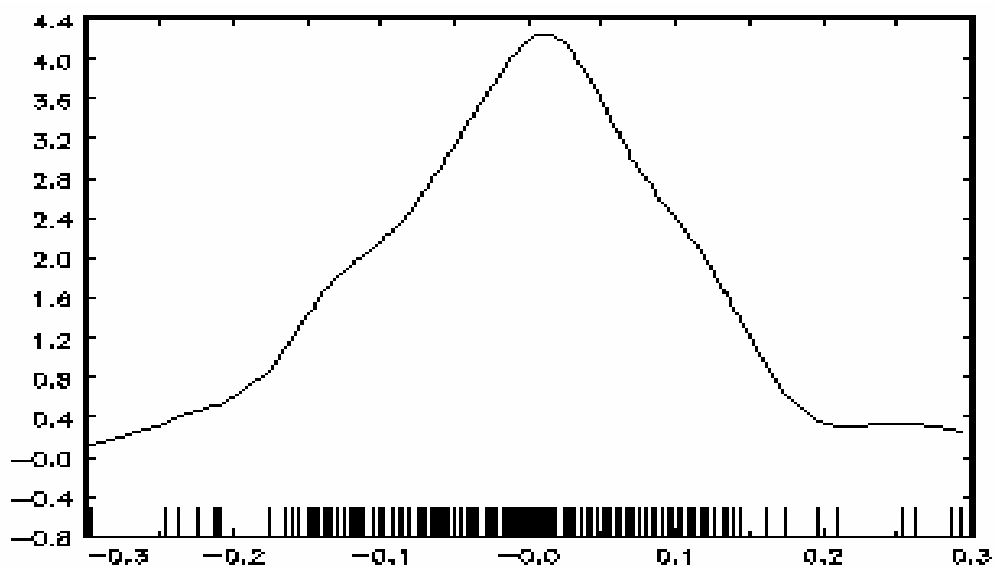
**Tabla 2:** Selección de modelo STAR

	Teräsvirta			Escribano-Jordá	
Variable Transición ( $s_t$ )	$H_{03}$	$H_{02}$	$H_{01}$	$H_{0L}$	$H_{0E}$
$\Delta_{12}y_{t-1}$	0.024	0.459	0.000	0.0464	0.229
$\Delta_{12}y_{t-2}$	0.038	0.744	0.000	0.0606	0.198
$\Delta_{12}y_{t-3}$	0.0826	0.869	0.001	0.0264	0.106

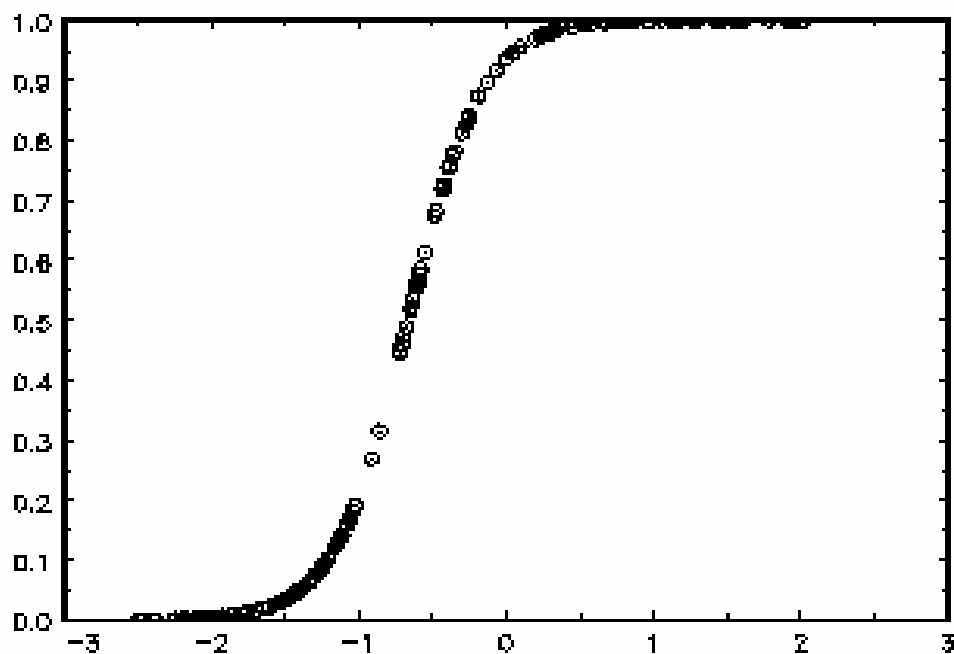
Nota: Los test son aplicados sobre un modelo AR(13) para la primera diferencia, incluyendo la variable en niveles retardada y las ficticias estacionales. En la especificación de Teräsvirta (1994), las hipótesis chequeadas son  $H_{03} : \beta_3 = 0, H_{02} : \beta_2 = 0/\beta_3 = 0, H_{01} : \beta_1 = 0/\beta_3 = \beta_2 = 0$  sobre (11). Si el p-value de  $H_{02}$  es inferior, un modelo ESTAR será el elegido, siendo un LSTAR para el resto de los casos. El procedimiento de Escribano-Jordá consiste en la comprobación de las hipótesis  $H_{0E} : \beta_2 = \beta_4 = 0$  y  $H_{0L} : \beta_1 = \beta_3 = 0$  sobre (17), siendo un modelo LSTAR (ESTAR) seleccionado si el mínimo p-value es obtenido para  $H_{0L}(H_{0E})$ .



**Figura 5:** Residuos del modelo LSTAR correspondiente a la tasa de desempleo mensual española



**Figura 6:** Función de distribución Kernel de los residuos del modelo LSTAR estimado



**Figura 7:** Función de transición del modelo LSTAR correspondiente a la tasa de desempleo mensual frente a la variable de transición  $\Delta_{12}y_{t-1}$

**Tabla 3:** Diagnostic Testing del Modelo LSTAR estimado

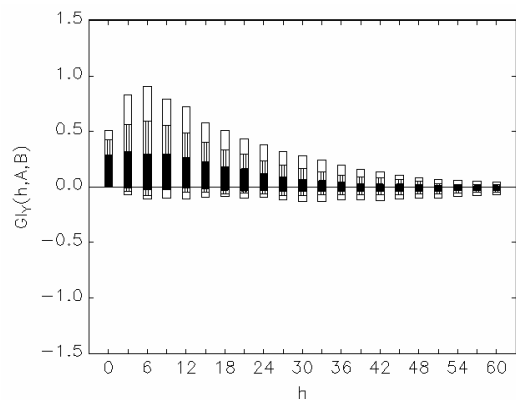
	Test de Correlación serial de orden $q$					
$q$	2	4	6	8	10	12
$p$ -value	0.965	0.488	0.791	0.686	0.750	0.652
	Test de Constancia de los parámetros					
	Constantes y Dummies			Variables dependientes retardadas		
	$LM_{C,1}$	$LM_{C,2}$	$LM_{C,3}$	$LM_{C,1}$	$LM_{C,2}$	$LM_{C,3}$
$p$ -value	0.139	0.089	0.122	0.149	0.138	0.359
	Test de no-linealidad restante					
	Constantes y Dummies			Variable dependiente retardada		
	$LM_{EMR,1}$	$LM_{EMR,2}$	$LM_{EMR,3}$	$LM_{EMR,1}$	$LM_{EMR,2}$	$LM_{EMR,3}$
$\Delta_{12}y_{t-1}$	0.218	0.689	0.710	0.609	0.579	0.627
$\Delta_{12}y_{t-2}$	0.314	0.638	0.630	0.669	0.575	0.723
$\Delta_{12}y_{t-3}$	0.388	0.661	0.603	0.662	0.693	0.641
$\Delta_{12}y_{t-4}$	0.478	0.726	0.619	0.640	0.668	0.685
$\Delta_{12}y_{t-5}$	0.629	0.777	0.466	0.425	0.563	0.663
$\Delta_{12}y_{t-6}$	0.560	0.734	0.674	0.492	0.711	0.569

**Tabla 4:** Evaluación de la predicción de los modelos AR y LSTAR

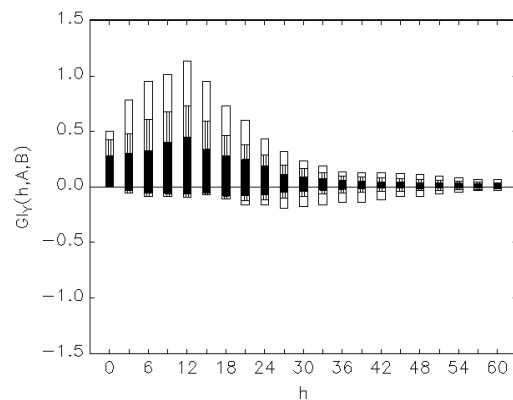
	MPE		MSPE	
$h$	AR	LSTAR	AR	LSTAR
1	-0.013	0.003	0.018	0.016
2	-0.005	0.004	0.018	0.047
3	-0.005	0.019	0.019	0.079
6	-0.008	0.038	0.018	0.132
9	-0.011	0.052	0.018	0.205
12	-0.009	0.089	0.020	0.166

Nota: El periodo de predicción se ha efectuado desde enero del 2000 hasta diciembre del 2002

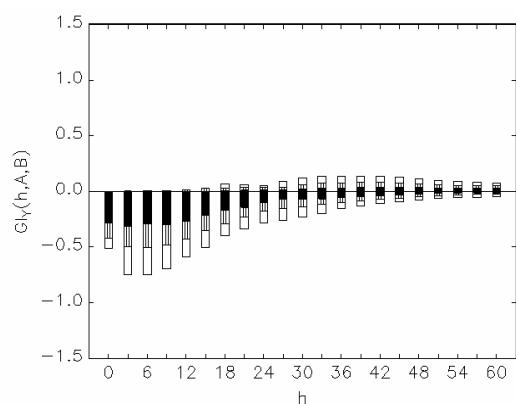




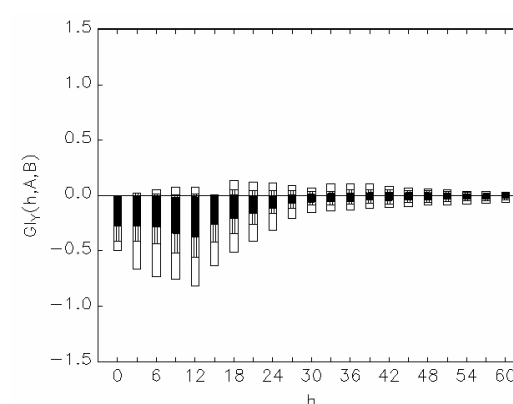
(a) Shocks positivos en recesión



(b) Shocks positivos en expansión



(c) Shocks negativos en recesión



(d) Shocks negativos en expansión

**Figura 8:** HDRs de las funciones de impulso respuesta del modelo LSTAR para el desempleo mensual español.

**Nota:** Los HDRs están calculados al 50, 75 y 90% (gráficamente en negro, gris y blanco, respectivamente). La recesión y expansión es considerada con diferentes historias, que tiene en cuenta los valores de la función de transición dependiendo de que  $G(\Delta_{12}y_{t-1}; \gamma, c)$  sea menor o mayor a 0.5 respectivamente.